EVALUATION OF EDUCATIONAL COMPUTER PROGRAMMES AS A CHANGE AGENT IN SCIENCE CLASSROOMS

Johnnie Wycliffe Frank Muwanga-Zake

B.Sc. (Makerere), M.Sc. (Rhodes),
P.G.C.E. (NUL), B.Ed., M.Ed., P.G.E. (Rhodes)

University of KwaZulu-Natal,

South Africa

EVALUATION OF EDUCATIONAL COMPUTER PROGRAMMES AS A CHANGE AGENT IN SCIENCE CLASSROOMS

By

Johnnie Wycliffe Frank Muwanga-Zake

B.Sc. (Makerere), M.Sc. (Rhodes),
P.G.C.E. (NUL), B.Ed., M.Ed., P.G.E. (Rhodes)

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy of Education in Instructional Design,

University of KwaZulu-Natal,

South Africa

December 2004

DECLARATION

I DECLARE THAT THIS IS MY OWN WORK, EXCEPT FOR THE SUPERVISION AND CITATIONS MADE IN IT, AND THAT IT HAS NOT BEEN SUBMITTED FOR ANY OTHER DEGREE AT ANY UNIVERSITY.

Date **December, 2004**

Signature

Name: Muwanga-Zake, Johnnie Wycliffe Frank

ACKNOWLEDGEMENTS

- Professor Alan Amory, University of KwaZulu-Natal – for a very constructive supervision and guidance that allowed my ideas to grow, and for the assistance in obtaining funding for the study.

- Professor Norman Thomson, University of Georgia (USA).

- The National Research Foundation of the Republic of South Africa for funding part of this study.

- My family, for support and patience.

ABSTRACT

This evaluation started with preliminary research into the situations and problems in science classrooms and computer laboratories. The preliminary research identified teacher-centred lessons, learner and teacher conceptualisations, large numbers of learners per classroom, assessment, and a lack of interest in biology as some of the major problems in South African classrooms. The current research (because it is continuing) uses two Educational Computer Programmes (ECPs); a Computer-Aided Assessment (CAA) programme which is designed to alleviate problems in assessment, and Zadarh (a constructivist adventure game) designed to solve problems in biology classrooms, to further investigate some of the identified problems and find out the learners' and teachers' views on the utility of these two ECPs. The use of these two ECPs had not previously been investigated appropriately, especially in disadvantaged communities where teachers had little knowledge of the use and of evaluating ECPs.

Therefore, a major concern for this study is that previous ECP evaluations excluded teachers and were not comprehensive enough especially for deploying ECPs in disadvantaged communities. A review of the methods that had hitherto been used, indicated that quantitative, mostly, behavioural and cognitive, pre-test post-test methods were prominently used, despite the shift in instructional design to constructional design, which embrace qualitative aspects of learning. Also, instructional design has evolved from behavioural models to include constructivist microworlds, which were unfairly evaluated by excluding qualitative benefits.

Thus, this study seeks a more comprehensive evaluation strategy, in which teachers play the role of co-evaluators and which captures the qualitative and quantitative changes that software programs impart upon teachers' classroom practices, with sensitivity to the multiple disciplines in a program, as well as to the value systems of teachers.

Comprehensive evaluation processes were facilitated during which 26 teachers in 23 schools in the Eastern Cape, KwaZulu Natal and Mpumalanga Provinces embarked upon the evaluation of the two ECPs. Evaluations were based upon a developmental, constructivist and interpretative approaches, by which teachers took ownership of

these evaluations.

Comprehensive evaluations revealed benefits from CAA and Zadarh, as well as benefits from direct teacher participations in the evaluations. CAA (Question Mark in this case) instantly provided diagnostic data. However, it was evident that the quality of diagnosis and remediation depended upon the quality of the test items, and the learning as well as the teaching strategies. Factors that could militate against the use and full utilisation of CAA in the schools where the study was done included the cost of software for CAA, teachers' capacity to set diagnostic test items particularly in a multiple-choice format, teachers' ability to interpret data produced by CAA, and teachers' skills in remedying their classroom problems as well as learners' problems.

This study found that by playing Zadarh learners were able to construct knowledge through discovery and were attracted to the enjoyable aspects of this educational tool. Learners remembered most of those moments in the game during which they were both stuck and trying to solve problems on their way through Zadarh. Therefore, Zadarh can provide useful learning experiences with fun, and can improve motivation towards learning.

Debilitating factors against the use of Zadarh and CAA include school curricula, which do not accommodate innovations, inflexible timetables, and classroom approaches that are teacher-centred.

It was clear that the success of using computers in education would depend upon the ability of teachers to evaluate the ECPs, and to integrate ECPs into school curricula. Building social networks with the teachers and the use of social constructivism to drive these interactions played an important role in the successful integration of ECPs into classroom. One way of achieving such success is to include teachers as evaluators and co-designers of ECPs. Evaluations of ECPs therefore should: i) allow the teachers and learners, through social dialog, to identify how software could solve problems; ii) establish the compatibility of the software with the school curriculum; iii) ascertain the capacity of school computers to execute the software; and iv) provide support to the teachers in the use the software. Evaluations should benefit teachers and learners.

The study concluded that a post-modern, developmental, and constructivist evaluation process might be one of the ways of enhancing training teachers in the use of the ECPs, in the concepts that the software deal with, and in evaluation. In that way, a socially contracted evaluation is comprehensive and can serve as a change agent through which teachers reflect and act upon improving their classroom practices.

CONTENTS PAGE

DISCUSSION 1

DISCUSSION 2

TABLES

FIGURES

APPENDIX

MANUALS

A BRIEF CLARIFICATION OF SOME KEY CONCEPTS

- **Curriculum 2005**

A reformed curriculum for South African schools introduce in Year 1998. Traditional subjects were integrated into Learning Areas. For example, Physical Science, Biology, and Geography formed Natural Sciences. It is a post-modern, constructivist curriculum that uses Outcomes-Based Education (OBE). Continuous Assessment complements traditional examinations.

- **Department of Education (DoE)**

DoE stands for the Department of Education of South Africa. The DoE sets policy on education, including assessment policy and the overall national curriculum. Policies can be obtained on the government website (http://education.pwv.gov.za/)

- **The National Research Foundation (NRF)**

This is a South African government body that identifies and finances research in South Africa.

- **Subject Advisor**

This is a title given to officials from the DoE who advise teachers about curriculum issues in South Africa.

- **Educational Computer Programme (ECP)**

An educational programme delivered by a computer

- **Disadvantaged and advantaged schools**

In this thesis, disadvantaged schools are those that had never used computers and obtained them through projects (i.e., could not afford computers). These were Black schools often in townships or rural areas. The advantaged schools were formerly 'Whites-only' schools and afforded buying computers without project intervention.

One school was disadvantaged and the other advantaged in the evaluation of CAA. While only 2 schools (out of 23) were advantaged in the evaluation of Zadarh.

PART I

INTRODUCTION

Several theorists and professionals (E.g., Ogunniyi, 1997, 2000; Manzini, 2000; McComas, Clough & Almazroa, 1998; Muwanga-Zake, 1998) identify numerous problems in science education in South Africa (SA) and computer technology is seen as one way of surmounting problems in science classrooms. However, many educators, including science teachers in South African secondary schools misevaluate the Educational Computer Programmes (ECPs). Therefore, schools in developing communities have recently become dumping grounds of ECPs.

Furthermore, some instructional designers and marketing agencies recommend computer programmes using data collected from pre-test post-test quantitative evaluations, and do not offer guidance on how teachers can incorporate the programmes into school curricula. Conventional evaluations of educational computer programme (hence forth ECP) rely upon simplified causal relations, which exclude most of the teachers' experiences (Peled, Peled, & Alexander, 1991: 419-448). These traditional evaluation models (such as applied in Young [1996], and Cates & Goodling [1997]) are not sensitive to judgements that end users make (Barbera, 2004, 13), and to qualitative learning attributes of ECPs. For example, written tests do not completely measure benefits obtained from computer games (Hogle, 1996). All of these factors might provide some explanation as to why previous evaluations have been inconclusive on the benefits accrued from using ECPs, and why such evaluations show incongruence between evaluator's conclusions and what teachers say (Randel, Morris, Wetzel, & Whitehill, 1992; Adams, 1998; Stratford, 1997).

RATIONALE FOR THIS STUDY

There is a need for an evaluation scheme that is considerate of the qualitative dimensions and of the disciplines of knowledge embedded within an ECP, and which involves teachers and learners in the evaluation. Such an evaluation checks on the teachers' evaluation competency and conceptual understanding of the subject matter, and would be sensitive to the realities of the school curriculum with respect to the use of ECPs.

The main concern for this research is to devise ways of evaluating ECPs more holistically, which check on a wider range of factors involved in implementing ECPs in a school, and on both quantitative and qualitative outcomes accrued from using ECPs, including attitudes, skills, quality of engagement with users, and curriculum compatibility (Peled *et al.*, 1991: 419-420; Randel, *et al.*, 1992, 261; Harlen, 1993: 60-66; Young, 1996: 17; Heinecke, Blasi, Milman & Washington, 1999: 1-2).

I used science ECPs towards developing such a holistic evaluation scheme because science is an important subject. For example, it is believed that a greater number of science graduates increases the capacity of production of goods, and in turn contributes towards development (Robottom & Hart, 1993: 591). However, there is a shortage of trained science graduates in South Africa (SA), and one possible cause of this shortage is the high failure rate in science at school (DoE, Executive Summary of the Report of C2005 Review Committee, 2000: 5). In trying to increase the number of science graduates, the Department of Education in SA (DoE) and the National research Foundation (NRF) have encouraged learners to take science, trained more science teachers, and have supported many projects on science and science education. Evaluating science ECPs such as CAA and Zadarh that can be used to improve science learning and teaching, especially in disadvantaged communities, contributes towards the efforts of the DoE and the NRF, since an increase in the number and quality of scientists *inter alia* requires SA to research the problems in science classrooms and find solutions to them.

RESEARCH QUESTIONS

The following were the main research questions:

1. **What does designing a comprehensive evaluation instrument for educational computer programmes in disadvantaged communities entail?**

2. **How does the teachers' participation in evaluating educational computer programmes influence the teachers' classroom practices?**

Three subsidiary research questions provided answers to this question using two ECPs, CAA and Zadarh, in a sample of schools in SA:

a. How well can computer-aided assessment (CAA) be used to provide teachers with diagnostic information on learners' current understanding of the science concepts at grade 10 level?

b. How does Zadarh help to solve some of teachers' problems in grade 10-12 biology classrooms?
c. What values do learners and teachers attach to CAA and to Zadarh?

TWO PROGRAMMES WERE EVALUATED

CAA was evaluated first, and provided insight into the evaluation process, which the teachers and I used in the evaluation of Zadarh. It was presumed that these two ECPs would be valuable to teachers and learners in the disadvantaged communities in SA.

1. Computer-Aided Assessment

What is Computer-Aided Assessment?

In simple terms, Computer-Aided Assessment (CAA) should mean the use of computers to aid assessment. However, CAA and other similar terms such as Computer-Assisted Assessment and Computer-Based Assessment (CBA) have been used variously, sometimes to mean the same thing, and other times used totally for different things. For example, although Wise & Plake (1990) refers to Computer-Based Testing (CBT), and others such as Sandals (1992), Thelwall (2000), Gretes & Green (2000) as well as Twomey & Miller (2000) refer to CBA or CAA, they do not provide its precise definition, but look at CBA as embracing all assessments in which computers are used.

Bojic (1995) is one of the few authors who attempts to differentiate between Computer Assisted and Computer Based assessment. Bojic reports that the term 'assisted', and the term 'aided, have been used interchangeably. The problem though is that Bojic reserves the term 'testing' for Computer Assisted, and assessment for Computer Based – i.e., Bojic writes of Computer Assisted/Aided Testing, and Computer Based Assessment. However, Bojic defines Computer Aided Testing as the use of technology *to manage or support the assessment process,* and defined Computer Based Assessment as doing assessment *via the computer.* To Bojic, assessment is computer based if the computer does all processes such as recording, feedback, and analyses.

Gretes & Green (2000: 47) state that Computer-Assisted Assessment is *a situation* in which computers are *used to administer and score practice tests on demand, thereby* *giving [learners] immediate feedback about their preparation for a paper-and-pencil exam to be taken during class.* That is, computers are only used to prepare learners for some final examination. But Thelwall (2000: 38-39) mentions that CBA is used for examinations and diagnostic assessment, particularly in mathematics and science.

In this study, I chose to use the term Computer-Aided Assessment (CAA) and define it as the use of computers for assessing learners' understanding, involving teachers setting and loading a test on a computer, the computer storing a data-base of questions, learners answering questions on the computer, as well as the computer marking, processing data, and providing a feedback on the assessment.

Furthermore, I think that the term 'aided' is more suitable because it signifies the fact that the computer is an aiding tool to other important forms of assessment that are not done on the computer.

Previous studies on CAA

CAA has received attention over the years. It is enough at this point to state its attributes that attract such attention.

According to Thelwall (2000), Croft, Danson, Dawson, & Ward (2001: 53), as well as Gathy, Denef, & Haumont (1991: 109), CAA, provides opportunities to learners for assessing themselves frequently at their own pace. Such multiple attempts, all of which are instantly analysed, can help teachers and learners to discover quickly and frequently how much the learners know, and what mistakes they make. CAA analyses each learner's responses in detail thereby supporting the diagnosis and remediation of a learner's problems. Thus, CAA was particularly evaluated for its potential to provide detailed diagnostic data for large numbers of learners. But, these advantages are mixed with disappointing findings, and more research is required.

There are many companies that produce CAA software. In this evaluation, CAA was set on Question mark (QM). The procedures for using QM are well documented by 'Question Mark', at http://www.qmark.com/. In this study, the teachers and I set and typed a test into QM *Designer,* version 1993. This QM Designer version can run on Windows 3.11 platform, as well as later versions of Windows.

QM designer allows a wide range of questions, including hot spot, word-match, type-in explanations, and multiple choices. Then QM Reporter analyses and makes various reports, such as individual performances and number of trials, class average, test performance, etc.

This study

I tried to answer the following specific questions on CAA:

How well can CAA be used to provide teachers with diagnostic information on learners' current understanding of the science concepts at Grade 10 level?

Answering this question required preliminary investigations. First, the QM 1993 version uses Multiple-Choice Questions (MCQs). It was therefore necessary to find out whether teachers could set diagnostic MCQs, by attempting to answer the following question: How diagnostic are the MCQs set by teachers on Grade 10 science concepts?

Second, if MCQs are to be accurately diagnostic, they have to be valid. Therefore, it was important to check on the validity of the MCQ items and to investigate the problems teachers had with diagnostic assessment.

2. Zadarh

Introduction

Zadarh is an adventure game and is clearly described by Amory (2001). I outline aspects that are relevant to this study.

The main objective of Zadarh is to get learners interested in studying biological concepts that literature has revealed as problematic to South African learners. Zadarh should also increase curiosity, as well as their problem-solving and critical thinking skills among learners, such that the broad focus is to support learners in gaining experiences of scientific inquiry, and if possible apply these situated experiences in other real life adventures.

Theoretically, Zadarh is based upon play as a 'human activity' comprising of a number of acts (as described by Newman & Lamming in Amory, 2001). These include exploring a visualisation space in which attempts to overcome the challenges helps the player or learner to discover the concepts that the game teaches. The explorations could involve cognitive (interpretive) or social constructions. Designers of Zadarh assume that exploration of the game involves fun and drama, which are accentuated through the interaction of a learner with the graphics in the game.

The game represents a model in which there is dialectic between pedagogical dimensions and game elements. Thus, Zadarh should, in the long term, enhance understanding of biological concepts, specifically selected on account of the evidence that learners find the topics difficult. The concepts Zadarh include photosynthesis, respiration, genetics, and evolution.

Research done on Zadarh before this study

Adams (1998) investigated the effectiveness of Zadarh in learning the relationship between respiration and photosynthesis at Matric level, and at two tertiary institutions using pre-testing and post-testing of participants who had computers at their homes.

According to Adams (1998), Zadarh anchored learning activities to authentic situations such as filling gas cylinders, which gave ownership of the learning process to the learners, and an opportunity of testing their alternative conceptions. Zadarh helped learners to gain a clearer understanding by challenging their previously held views by solving puzzles in Zadarh, while at the same time contributing to enjoyment. Participants believed that playing computer games depended on previous experiences and the ability to interpret the user interface (Adams, 1998: 64). Furthermore, Zadarh helped learners to collaborate during play, and to enhance recreation.

Adams found that realistic goals, fantasy, mystery, solving puzzles, navigation skills, attitudes, and adequate feedback (including the associated sound) contributed to the enjoyment of Zadarh, but without competition. With regard to design, Adams (1998) found that learners noted the limitation of mobility, but felt that navigation became easier with experience.

Learners felt that there were enough clues to solve puzzles, but recommended more guidance, a defined start and finish of the game, provision of a floor plan, voice instruction, more furniture, different kinds of music, and more feedback. It would be better to save the game so that one can start where he stopped. Adams (1998: 81) recommended that Zadarh might be used as both a learning and assessment tool in the presence of a facilitator. Adams concluded that Zadarh did not help learners to overcome deep-seated misconceptions related to respiration and photosynthesis.

Additional research by Ivala (1998) found that Zadarh improved collaboration among learners, and also improved their understanding of respiration and photosynthesis. Ivala concluded that realistic goals, fantasy, mystery and adequate feedback contributed towards enjoyment of computer games while learning.

This study

This study was interested in the nature and quality of playing Zadarh, and how this could support learning biology. These interests were investigated through answering the following question: Does Zadarh help to solve some of teachers' and learners' problems in Grade 10-12 biology classrooms?

Therefore, I had to conduct a preliminary survey and investigation into the problems biology teachers face in Grade 10-12 by answering the following questions:

- What are the teachers' problems in biology classrooms?
- How best can Zadarh complement the teaching methods employed by the teachers?

TWO LEVELS OF EVALUATION

I conducted two levels of evaluation: The first level of evaluation was *validation* (Percival & Ellington, 1984: 119; Mackay, 1975: 194; Greene, 1994: 531) for CAA (generally) using a version of Question Mark (QM), as well as of Zadarh. By preliminary and literature surveys, validation was a process of establishing whether the CAA and Zadarh were suitable for the context (the Grade 10-12 science schools in disadvantaged communities in South Africa), the readiness of potential end-users (the teachers) to use the programmes, and the nature of the science curriculum in SA (Newman & Lamming, as cited in Amory, 2001). This first evaluation level also looked at how the two ECPs could fit into the science (biology and physical science) curricula by investigating the way teachers taught.

Validation included establishing whether teachers needed the two ECPs. The purpose of a needs assessment is to see how a learning programme serves important needs of specific audiences (Reeves & Hedberg, 2003: 119; Shakeshaft, 1999: 3). I had to find out whether CAA addressed problems in assessment, and Zadarh addressed the lack of interest or motivation for biology among learners, and upon the teachers and learners' poor conceptual understanding of biology.

Another important dimension investigated was the capacity of teachers to use, and computers to run the two ECPs.

The second level of this evaluation and the main activity of evaluation was to find out the values teachers and learners attached to CAA and Zadarh after they had experienced these two ECPs in their schools. Teachers, learners, and I judged these experiences against the usual classroom forms of delivery of assessment and of learning. That is, these judgements described the *values* (Lloyd-Jones, Bray, Johnson & Currie, 1986: 1; Fink, 1995: 10; Weiss, 1998: 5; Mackay, 1975: 190 - 191) with regard to assessment and learning.

THE FIRST LEVEL OF EVALUATION – IDENTIFYING PROBLEMS IN CLASSROOMS

Introduction

I motivate the need for interventions in assessment and in biology education by outlining the problems in this section. Nonetheless, I have to point out that some ECPs are not inspired by problems but by interest, innovation, and possibly, intuition. The following were the sources of information.

The original sources of information were my experiences of teaching physical science and biology for over 20 years at different levels in SA. I had also done research on biology before my teaching career, and more research on problems in science classrooms at a Masters degree level in SA before conducting this preliminary survey.

In the preliminary survey from 1997 up to 2001, I investigated the problems in assessment with 2 Grade 10 teachers and 27 learners in East London (Eastern Cape province) in 1998 and 1999. I also researched problems in biology classrooms with many teachers, but I present findings from 26 Grade 11 – 12 teachers and 192 learners in 23 schools in the Eastern Cape, KwaZulu Natal, and Mpumalanga provinces during the third and fourth quarters in 2002 and 2003. I have also been discussing the problems during teacher science workshops in the course of my work from 1997 to 2004 at the Centre for the Advancement of Science and Mathematics Education (CASME).

What problems do teachers face in assessing science using diagnostic Multiple-Choice Questions (MCQs)?

Why focus on diagnostic MCQs?

We use MCQs extensively in assessing science in SA. For example, a science Matric examination can comprise 25% MCQ worth of marks. I have experienced problems in setting good quality MCQs in my teaching career, and as a Matric examiner. I have also found that the quality of MCQs set by the teachers I supervised as a Head of Division of science at schools in SA was often poor.

In this preliminary study, I identified problems teachers faced in setting science MCQs for diagnostic purposes by asking science teachers to set a Grade 10 science test. Then the teachers and I revised the questions to make them diagnostic. An Instructional Designer, a Science Subject Adviser for East London schools, two teachers, and later two focus groups of learners validated the test items. The aim of this preliminary survey was to find out whether teachers needed CAA by providing answers to the following question: What problems do teachers face in assessing science using diagnostic MCQs?

Teachers' problems in setting diagnostic MCQs

The science test teachers set appear in Test 1 (see Appendix I), but it comprised recall (knowledge) items mainly, and some of the statements were not very clear. I had to guide the teachers to include higher order cognitive processes, to improve the stems, to improve upon distracters, and to include different kinds of multiple-choice items such as 'matching items'. The other problem was that teachers had difficulty in setting and in answering diagnostic items, apparently because of the problems they had in conceptual understanding of science.

Problems in diagnosing learners' problems and remedying their problems

Besides the teachers' poor conceptual understanding of science, large numbers of learners per class inhibited the prospects of teachers' use of diagnostic assessment. Diagnosis requires an analysis of answers of each individual learner, which takes too much time in a class of, for example, 60 learners. CAA could help in speeding up the analysis of individual learner's responses.

Conclusion – problems Grade 10 South African teachers face in assessing science by diagnostic MCQs

The teachers' problems include setting good quality MCQs, setting diagnostic MCQs, and using the results to diagnose learners' problems. These problems relate to conceptual understanding, and most likely show that these teachers did not receive quality science education and teacher training. The other problem is the normally large numbers of learners per teacher, which limits the time a teacher can spend attending to a learner.

What are the teachers' problems in Biology classrooms?

Why focus on biology?

Biology attempts to explain the functions of vital life processes and phenomena (Nagel, 1953: 537-540; Schlick, 1925: 523). Epidemics such as HIV, cholera, and TB, as well as debates on human and livestock health, genetically modified food, and environmental degradation, require a public well equipped with biological knowledge. Furthermore, most Grade 12 learners study biology among science subjects (Matabane, as cited in Chacko, 1996), because they think it is an easy subject.

The biology curriculum in South African schools

Biology is taught at school as 'Biology', or as part of General Science, and in Curriculum 2005, as part of Natural Sciences, under the theme 'Life & Life Processes'. General Science in Grade 9 comprises use of apparatus, materials and other aids in the study of biology; the cell; morphology and functions of the parts of flowering plants; sexual reproduction in angiosperms; life processes and systems of humans. Biology syllabi in Grade 10 includes ecology, the cell, cell division, plant tissues, angiosperm anatomy, mammalian tissues, aspects of the anatomy and of the physiology of humans. In Grade 11, learners study viruses, bacteria, some plant types, some invertebrate types, cell division and genetics, as well as some aspects of the anatomy and physiology of humans. Then, Grade 12 learners study biological compounds and nutrients, enzymes and co-enzymes, angiosperm physiology, cellular respiration, some aspects of human anatomy and physiology, some aspects of homeostasis in certain animals and humans, as well as some aspects of population dynamics. Biology can be studied at Standard Grade (SG) or Higher Grade (HG), although C2005 is doing away with SG and HG. Learners are expected to apply knowledge to new situations, which may be unfamiliar and to do practical work.

Problems in biology classrooms

Indicators of problems in biology education

a. Poor Matric results

Chacko (1996) reported a high failure rate especially after the introduction of a new syllabus. The poorer biology Matric examination results indicate the possibility that biology is in a worse situation than physical science (Table 1).

Table 1: National South African pass rates at Matric (Minister of Education, 2003)

SUBJECT	GRADE	2001	2002
PHYSICAL SCIENCE	HG	72.4%	76.3%
	SG	66.9%	76.4%
BIOLOGY	HG	58.7%	65.3%
	SG	72.1%	74.4%

b. Examiners' reports

The DoE reports of 1997 (29-33), 2001(41), and 2003 (43- 49) for KwaZulu-Natal province show that certain problems have persisted since 1997. These include:

- Language; mainly that learners did not understand instructions and some of the statements used in questions (E.g., differentiate, tabulate, explain, describe). Learners misunderstood 'biology language' and wrote meaningless answers that include irrelevant information and a lot of spelling mistakes. Other sources confirm this problem (Sanders & Mogodi, 1998; KwaZulu Natal Provincial Matric Examinations Report, 1997: 29).
- Conceptual understanding; this was manifested through misconceptions (E.g., in population dynamics and enzymes), failure to differentiate related concepts (E.g., auxins/axons; tissue fluid/lymph; excretion/secretion), and learners' lack of critical insight. Learners mainly answered content questions, and fail to relate function and structure.
- Inadequate or uncompleted practical work; learners' answers showed that they never complete or do practical work. Therefore, learners did not demonstrate process skills, application of information, and interpretation of

data. Graphing is not well done. (E.g., candidates could not plot on the X and Y axes, and assumed that all graphs referred to population dynamics). It also appears that learners memorise some procedures.

- Learners did not know how to draw and to label diagrams
- Indications that the syllabus was not completed (E.g., many questions not answered)
- Learners were not able to recall knowledge from lower grades: E.g. Grade 10

c. Opinions of a Regional Subject Advisor for Biology in KwaZulu Natal

A discussion with a Regional Subject Advisor for biology in KwaZulu-Natal province during September, 2002 revealed the following:

> 'At Matric, learners found Plant-Water relationships, the Nervous Co-ordination, and Excretion most difficult. Naidoo explained that learners did not seem to capture well the experiments, and lacked the conceptual understanding involved in those experiments. With excretion, the processes in the kidney, such as osmo-regulation, and the role of hormones were not understood'.

He also pointed out that Standard Grade learners did worse than Higher Grade learners, and that

> 'At Grade 11, plant biology was done badly because it seemed to be boring to learners and teachers. Learners also lacked an understanding of the sequences in respiration and photosynthesis. The reason for this appeared to be a watered down syllabus that missed out important facts. Learners did not understand crossing, and left out possible genotypes'.

I also learnt from the *Subject Advisor* that the DoE does not monitor the performance per Biology topic to the extent that one would have to visit the schools and find out.

d. A drop in the number of biology students

Introduction

There is evidence that fewer learners are opting for Biology (Muwanga-Zake, 2000). Furthermore, the DoE in some provinces, such as the Eastern Cape, persuaded learners to opt for the easier Standard Grade Biology to increase the pass rates. Pass rates fell further, indicating that performance in biology was falling.

Teacher complaints

Reports on South African problems in biology over the years consistently report that teachers' complaints included long syllabi, terminology problems, useless objectives, illogical sequencing of topics in the syllabus, examinations that test for facts, and a curriculum development process that does not recognise teacher inputs (Sanders, 1995: 722; Wood-Robinson, Lewis, Leach, & Driver, 1997; Sanders & Mogodi, 1998; Ivala, 1998; Sanders, 2002: 85). My survey involving the 26 Grade 10-12 teachers was consistent with some of those reports as follows:

i. Syllabus too long, especially Grade 12 (57%);
ii. Grade 11 syllabus not related to Grade 12 syllabus (no continuity), and lack of logical arrangement of topics (45%);
iii. Lack of resources to perform practical work (34%);
iv. Difficult concepts (28%);

Insignificant numbers of teachers identified these as problems: abstract concepts; lack of textbooks (9%); data need to be analysed and learnt; difficult practical exercises; lack of motivation; poor study habits, and; lack of relevance to real life (each 7%).

In agreement with the examiners' reports, teachers pointed out that learners suffered from: confusion between related concepts (E.g., gaseous exchange and cellular respiration; poor interpretation of questions and data; relying on memorisation; neglect of practical work; lengthy experiments; confusion between structure and function; poor training of teachers, and; under qualified teachers. With regard to content, the 26 Grade 10-12 teachers in the survey confirmed Ivala's (1998) finding that there were problems in learners' understanding of population dynamics, respiration, genetics, diagrams, and molecular genetics.

Teacher qualifications

The majority (63%) of teachers had teaching diplomas and that the highest level they had studied biology was at that diploma level (Table 2). None had experience of teaching beyond 16 years, and about half of them (53%) had experience of up to 5 years of teaching.

Table 2: A sample of qualifications of Grade 10-12 biology teachers (n=26)

Teaching qualification	None	Certificate	Diploma (E.g., STD)	Degree/HDE/ACE
			63%	***37%***
Highest level studying *Biology*	Matric	Certificate	Diploma	Degree
	22%		***52%***	***26%***
Number of years teaching Biology	0- 5	6-10	11-15	16 and over
	53%	***33%***	***14%***	

Conclusion – problems in Grade 11-12 South African biology classrooms

Findings concur with most of the problems cited by Sanders (1995; 2002) and Sanders & Mogodi (1998), as well as the concerns examiners raised about the biology syllabus. The most popular reasons the 26 Grade 10-12 teachers gave for the difficulty of biology relate to syllabi lengths (57%) and illogical sequencing of topics (45%).

The 26 teachers also agree with Ivala's (1998) findings and the biology examiners' reports about the difficulty learners have understanding population dynamics, genetics, and respiration. Teachers presented this problem as 'difficult concepts' (28%) and 'abstract concepts' (17%). The later problem is concomitant with the scarcity of practical activities observed in classrooms, which in turn relate to the teachers' complaint of lack of science equipment, poor teacher training, as well as lengthy syllabi to be covered in limited time. Biology periods were shorter than an hour in 29% of the schools sampled. Lack of practical work might account for the learner's poor performance in questions that include data analysis.

Some of the difficulties teachers find with practical work might be related to their low academic standards, considering that 74% of them studied biology up to a diploma level. Diploma study does not exceed Matric level by far. Such levels do not impart conceptual understanding to a level that can innovate or critically look at practical activities. Hence, teachers complain of lack of resources for practical work even where some equipment would sufficiently work, and rely on textbooks, which they follow to the letter such that they also look for equipment that the book recommends. However, teachers did not admit to this apparent incapacity directly, but suggest that they would improve by further workshops. Thus, the source of poorly trained teachers is a vicious cycle, started by the majority of these poorly trained teachers, whose skills are recycled through learners.

What methods do teachers use to teach?

Time spent on biology

Forty two per cent of teachers engage over 51 learners in a class, while only 33% have classes of 40 or less. Biology lessons are long (27% are over 55 minutes), and 46% have 46-50 minute biology periods in a day (Table 3).

Table 3: Number of learners, and time spent on Biology (n= 23 schools)

Number of *Biology* learners/class (= Total number / number of classes)	*0-20*	21-30	31-40	41-50	> 51
	08%	***08***	***17%***	***25%***	***42%***
Duration of a period (Minutes)	25-34	35-44	45-54	*55-64*	*Over 65*
		16%	***13%***	***67%***	***4%***
Average time for *Biology*/class/day (Minutes) = Duration of a period x number of periods/5 days or days in a cycle	35-40	41-45	46-50	51-55	56-60
	19%	***8%***	***46%***		***27%***

Class observations

I followed the following procedure to obtain these records:

In KwaZulu Natal, SA, during September 2003, I had to obtain permission to observe and record science classrooms. With permission of the principal and teacher, I sat at the back of a class with a blank table (Table 4). Then I placed a tick in the relevant box when I observed an activity. The activities appear in the first column of the table and represent some aspects of learning and teaching strategies, especially for science. I wrote on a separate paper when an activity was not represented in the first column of the table. The first column reports teachers' and learners' activities in the class, while the first row indicates the time as the lessons progressed. Data has been combined for seven classrooms. The numbers (coefficients against symbols) inside the table indicate the number of occurrences of a particular activity.

The outstanding patterns in Table 4 are:

Teachers stated facts throughout the lessons, and gave procedures for practical activities after about 12 minutes (row 1). When teachers were not talking, learners (row 6) were doing practical work, but learners rarely contributed ideas (row 3). Row 8 confirms learners' participation in the lesson: although learners were involved by use of their psychomotor skills, they did not solve practical problems. The teachers' questions were mainly recall (row 4).

Practical work

Eighty eight per cent of the teachers believed that their learners enjoy practical tasks. However, class observations (Table 4) indicate that practical exercises did not involve learners' ideas, and so practical work was teacher-centred. Secondly, a problem with Table 4 is that the number of teachers who conducted practical work is obscure (row 6). However, from my records, practical work happened in 3 of the seven classrooms. Row 6 shows that practical work was taking place after 24 minutes in 3 of those five conducted. That is, it would seem like practical work is popularly done after about 24 minutes.

Knowledge about practical work contributes only about 40% of total marks. Hence, one can pass the Matric examination by memorising theory. The few experiments attempted oversimplify concepts such as respiration, and curricula have no time for lengthy processes such as genetic assortments. Furthermore, it is illegal to experiment upon human beings and animals in light of the advent of HIV and AIDS and laws against animal abuse. Hence, biology teachers see the biology curriculum as too theoretical, irrelevant, textbook-based, and teacher-centred (Hart & Robottom, 1990), and learners learn biology by rote (Ivala, 1998; Smith & Simmons, 1992).

Table 4: Science lesson observation schedule – 7 different classes on different topics in KwaZulu Natal during September 2003

ACTIVITY OBSERVED	SEQUENCE OF EVENTS – 2 MINUTE INTERVALS (Just place the 'first letter (s)' of the activity observed)																													
	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60
1. Teacher statements/questions Facts or closed (F) Investigative problems (P) Procedures (Pc) Open-ended (O)	8 F	5 F	6 F/o	3 F/o	5 F	2 F/P	2 F/Pc	F	F/Pc	P/Pc	F	F	F	F		F	F	F/P	F/P									Pc	Pc	
2. What do learners do with teacher's ideas? Question them (Q) Discuss / modify them (D)							Q	Q/D	Q/D	D		O																		
3. Learners' Ideas?					1	1	1						1	1	1			1	1											
4. Questioning techniques Recall, repetition, factual answer (R) Elaboration, justification or explanation from learners (E) 3. Open-ended (O)	6 R	6 R/O	7 R	R/E	2 R/2 O/E	R/2 O	R/O	O	O							R	R	E		O		R			R					
5. The teacher and learners' ideas ? Question and discuss them (Q) Modify them (D) Put them into practical work (W)				W															Q	Q	W	W	D	D						
6. Practical work				1	1	1	2	1	2	1	2	3	3	2	2	2	2	2	1	1	2	1	1	1	1					
7. Does the teacher initiate all activities? (√)		1	1	1	1	1	1	2	1	1	1	1	1					1	1											
8. Skills among learners? Observation (O) Psychomotor (Ps) Problem solving (Pr) Hypothesising (H)			O O	O O	2 Ps/O	3 Ps/2 O	3 O	Ps/3 O	Ps/4 O	Ps/4 O	2 Ps/2 O	Ps/3 O	Ps/2 O	Ps/3 O	Ps/2 O	Ps/O	Ps/O	Ps/2 O	Ps/2 O	Ps/2 O	Ps/O/H	Ps/H	Ps/H	2 H	2 H	2 H				
9. Problems with the kits / Programme (√)																														
10. Assistance – Between learners (Bl) Assistance - Teacher-learner (Tl)					Tl				2 Bl/Tl	2 Bl/Tl	2 Bl/Tl	Bl	Bl	Bl	Bl	Bl	Bl	Bl	Bl	Bl	Bl	Bl	Bl	Bl	Bl					
11 Group work (GW) Individual work (IW) Teacher demonstration (TD)	2IW	2IW	2IW/Td	2IW	2IW	2IW	2IW	2IW	2IW/GW	2IW/Td	2IW	3IW	3IW/GW	IW/GW	IW/GW	IW/GW	IW/GW	IW/GW	IW	Td	Td/GW	GW	IW/GW	IW/GW	IW/GW	IW/GW	IW	IW	IW	IW
12. Enjoyment (E)			3	2	3	3	3	2	2	2	2	2	2	2	2	2	2	2	1		2	2	3	3	3	1				

These observations are supportive of Ivala's (1998) findings with twenty high school teachers from 19 schools in KwaZulu Natal, which showed that the majority of the teachers relied heavily on behaviourist theories (instructivist), and the chalkboard was the most widely used teaching resource, while providing extrinsic motivation.

Teachers' opinions about what their learners enjoyed

All teachers indicated that their learners enjoy group discussions. Eight-two per cent of the 26 teachers believed that learners enjoy taking notes much more than anything else (Table 5).

Table 5: Teachers' opinions on what activities their learners enjoy (n=26)

	Do not enjoy	**Enjoy**	**Enjoy very much**
Problem solving	47%	29%	24%
Taking notes	18%	41%	41%
Asking questions	25%	50%	25%
Talking about their experiences	24%	41%	35%
Group discussions		64%	36%

The relationships between the 26 teachers' opinions and 192 learners' opinions (from the same 23 schools) were mostly weak. The analysis of the Pearson's correlation processed by use of SPSS indicates that few of the observed relationships were very strong (Table 6).

Table 6: Pearson Correlations between difficulty and enjoyment

	Teacher opinion of most difficult to learners	**Learner opinion of most difficult topic**	**Learner opinion of least enjoyed topic**	**Least enjoyed topic to teach**	**Most difficult topic to teach**
Teacher opinion of learners' most difficult topic		0.67(**)	0.25	0.15	0.39(*)
Learner opinion of most difficult topic	0.67(**)		0.55(**)	0.18	0.21
Learner opinion of least enjoyed topic	0.27	0.55(**)		-0.15	-0.04
Least enjoyed topic to teach	0.17	0.17	-0.15		0.45(*)
Most difficult topic to teach	0.39(*)	0.21	-0.04	0.45(*)	

** Correlation is significant at the 0.01 level (2-tailed);

* Correlation is significant at the 0.05 level (2-tailed).

Note: i. 26 Teachers and 192 learners were involved. Therefore, correlation figures and levels of significance should cautiously be considered.

ii. Table 6 is presented the way SPSS produced it.

The strongest relationship was between 'Learner opinions of most difficult topic' and 'Teacher opinion of learners' most difficult topics (r = 0.67), which indicate that approximately 49% of teachers and learners agreed on the difficult topics. The second strongest correlation was that between 'learner opinion of difficult topics' and 'learner opinion of least enjoyed topic' (r = 0.55). This indicates that learners do not enjoy studying topics they find difficult. It can also mean that topics are difficult because learners do not enjoy them. Similarly, the difficulty teachers find in teaching was significantly and positively correlated with the boredom they found in teaching those topics (r = 0.45). Teachers also thought that the topics they found difficult to teach were also difficult to learners (r = 0.39).

Teachers' suggestions of what can make biology easy (and enjoyable)

Only one teacher (out of 26) believed that biology is easy because it is about real life issues. Teachers also provided opinions on how biology can be made easier for learners:

> Dramatising; Applying concepts; Use practical work; Apply own-pace learning; Use videos; Teach little work at a time; Use group work and discussions; Arrange excursions; and Use computers.

The suggestions teachers gave for making biology enjoyable were similar to those they gave for making biology easy. This showed a possible confusion that easy and enjoyable meant the same thing or that something is enjoyable if it is easy.

Additional problems in science classrooms

I have been exposed to additional observations in the course of my work at CASME, and have found these important problems in science classrooms. The following problems, relevant to this study, have been consistent, and were confirmed further in a survey and discussion with 104 Grade 10-12 science teachers who attended a Carnegie workshop during June 2004.

A poor quality of teachers

It is rare to find a science graduate among teachers. For example, there were only six among 104 science teachers in the Carnegie project. Furthermore, 98 out of the 104 teachers had studied science at the Standard Grade (SG), which does not go deeper in conceptual understanding, and excludes most of the practical work. One possible reason for this is that learners who do well at Matric end up in courses such as Engineering. These SG products are unlikely to be comfortable teaching HG (and might advise learners to register for SG. Thus, teachers' level of science is rarely at a standard that can allow an open constructivist approach, which gives learners a chance to explore science freely. This low level of science could account for the teachers' verbatim reliance upon textbook notes and practical instructions (Muwanga-Zake, 1998), the practice of chalkboard teaching (Jennings & Everett, 1996), the teachers' inability to use equipment that is not familiar, and learners' shunning of science.

Teachers' misconception of their problems

The teachers who diligently attend workshops, like the 104 teachers, see a need to improve their practices in class, and worry about the majority who do not obtain further training. However, these teachers seemed to prioritise the wrong needs, and to be unaware of other crucial problems. An example was their demand for science equipment, which they hardly use. Only 10% of the 104 teachers claimed to have done all recommended practical exercises either during their schooling or with their learners. In most schools, one would find most of the equipment badly managed, gathering dust, broken or neatly stored away, some expired chemicals some of which teachers cannot not identify, lack of inventories of equipment and chemicals, and in some cases, teachers who do not know the use of some of the equipment. Teachers believed in practical work, despite not having had practical experiences in their academic and teaching careers. It is possible that the ECPs could end up in a similar way – stored away, damaged, under-utilised, mismanaged, etc.

Status of computers and teachers' opinions about the use of computers in a sample of Dinaledi schools

Introduction

A South African government project named 'Dinaledi' issued 20 computers to each one of 101 schools with the primary aim of improving mathematics and science examination results. I conducted my research in some of these schools, besides other non-Dinaledi schools. Computers were new in most of these schools, although some schools did not have electricity or secure computer laboratories. Two of the 23 schools I worked with had lost some of their computers to thieves. Furthermore, schools were still grappling with how to use the computers.

Bybee & Ellis (1988: 159) advise *serious consideration should be given to teacher education attendant to the implementation of information technologies in schools.* This consideration is important for this study because computer use in schools might significantly influence the way teachers and learners use and evaluate ECPs.

Affordability of computers

All these schools complained of the limited number of computers. For example, the 20 Dinaledi computers could not serve effectively a school of up to 600 learners. Computers were still regarded a luxury to many schools.

Principals in the participating schools indicated that they had to get money from the schools' fund for maintaining the computers. Some schools have computer laboratory managers appointed from the school staff, but the majority use mathematics and/or science teachers to manage the computers.

Use of computers in schools

Teachers complained that authorities prioritised teaching computer skills to learners, although Dinaledi supplied computers for use in teaching science and mathematics. Four of the 23 schools, which participated in this research concentrated on teaching basic computer principles, and word processing. Three of those 4 schools additionally offered computer science as a Matric-examined subject.

Computer use in learning areas depended upon what programmes a school had bought. For example, 3 schools had bought physical science and mathematics programmes from 'Power Education', a South African company specialising in selling ECPs to schools. However, no school had programme evaluators and schools had just bought ECPs without evaluating them. Furthermore, teachers did not have an idea of how to incorporate such programmes in the school curriculum, except for revision. Hence, one teacher felt cheated. Apparently, the way schools utilise computers has not progressed since 1996, because The Education Policy Unit, University of Western Cape (2000) obtained similar findings during 1996-97.

A few schools can afford using E-mail and the Internet. Internet use is limited by the shortage of expertise, funds for paying the service providers and for the telephone line, and is often for the principal. I found 7 schools (out of the 23) with working e-mail and Internet. TELKOM had disconnected many schools from the Internet.

Schools are struggling to get their school Local Area Network (LAN) to work because schools do not have people to manage LANs. I found only 2 schools out of 23 where individual teachers, out of curiosity and interest, had set up the LANs in a working order. TELKOM had set up LANs in all the Dinaledi schools but did not train teachers on managing these LANs. So, LANs do not function in the 21 schools I visited either because the teachers do not know how to use the LAN or because the LAN has malfunctioned without repair. Interestingly, TELKOM set up the LANs in such a way that no one else other than TELKOM staff can re-set or repair them. Schools are under contractual obligation not to allow non-TELKOM staff to repair the LAN or the computers.

Teachers' qualifications in computer skills

Forty five per cent of teachers in sampled schools had never been trained in using computers, and all teachers wanted training (Table 7). Principals complained of the scarcity of personnel in schools and community who are skilled in using computers.

Table 7: Teacher's qualifications in using computers

Never	Trained myself	Trained at an institution	Trained through a project	I am getting training now	Not interested	I want o be trained
45%	14%	36%	5%	16%	0%	100%

Teachers' experiences

Thirty three per cent of teachers had never used computers to teach, and 53% had less than years' experience of using computers in education (Table 8). Only 13 % had experience of over 5 years.

Table 8: Teachers' experiences in using computers

Never	Less than a year	Up to 5 years	6-10 years	11-15 years	Over 16 years
33%	53%	10%	0%	03%	0%

However, it has to be borne in mind that these experiences comprised simple computer skills such as processing marks using spread-sheets (Excel), accessing the Internet for notes and e-mail, and word processing of tests.

Teachers' opinions about using computers in education

Analyses of the dialogues indicated that teachers construed computers as modern, which will impart cultures of teaching and learning, as well as provide an environment in which learners can test and develop their natural talents. Teachers particularly believed in the usefulness of the Internet, especially with regard to communication and research.

Other popular beliefs were that computers would help to provide information for research through the Internet (E.g., 'Obtains information from different sources'; 'It has up-to-date information'; 'Exposes one to the world through the internet', and; 'Can be good for research'). Related to this opinion, was the view that computers store useful information (E.g., 'Useful for retaining information (including school records'; 'Complements teachers' explanation'; 'Computers can store a lot of information'; 'Can record teacher so that learners can listen to the lesson at a latter date'; 'For revision', and; Stored materials can be re-used).

Other less prominent opinions among teachers, which are popular in literature (for example, Heinecke *et al.*, 1999) included enhancing basic learning skills, as well as promoting psychomotor skills, enhancing studying at own pace, processing information, use in assessment, and use in simulating reality.

Imagination about problems that computers can cause in schools

Teachers thought that computers might cause the following problems:

- Can affect our jobs
- Computers are too expensive and have to be bought for schools
- Teachers and learners must be trained to use computers
- There would be problems with large numbers of student
- Time consuming
- Computers attract thieves

Some teachers had some reservations regarding the possibility of creating learners who are lazy to think or do research.

Teacher confidence

The attitudes of teachers towards using computers in education were in general positive. Many of them indicated that they had access to professional training in which computers were used, especially those that are conducted by the Department of Education and School Net.

Sensitivity about computers

It is apparent that schools are likely to be extra sensitive and very careful with their computers. An evaluation of an ECP in such schools has to be considerate of a number of issues some of which are:

- A need to negotiate with the school authorities the date, time of the day, and duration of the evaluation exercise bearing in mind the sensitivity and sentiments with which schools keep their computers. Schools are still not sure about who can have access to computers, and any computer use in some schools is a special occasion.
- Teachers have to be trained in basic computer skills, running programmes, using programmes for teaching, and then evaluating those programmes.
- In such contexts, it is likely that teachers might conflate the evaluation of the computer as a tool in education as well as the problems they are facing with computers, with the evaluation of an educational computer programme. That is, the computer is exciting to first users but problematic to maintain in disadvantaged communities. Teachers might not differentiate between the computer and the programme it delivers.

(I adopt Hannafin's (1999) definition of a tool – *a means through which individuals interact with resources and act upon their own thinking*).

Conclusion – status of computers in the schools where the evaluations were done

Schools were struggling to maintain computers and to get qualified teachers to run the computer laboratories. Some schools had bought ECPs they could not effectively use. Furthermore, computers were still new to many teachers.

Conclusion from the first level of evaluation: Validating CAA and Zadarh

On the basis of the problems in the above, the following are among the serious needs for Grade 10-12 science teachers in SA, which the two ECPs (CAA and Zadarh) could address:

- Learning how to set quality and diagnostic MCQs
- Ability to diagnose problems in conceptual understanding of large numbers of learners
- Improving upon learners' enrolment and interest in biology

However, these two ECPs had not been evaluated in these disadvantaged communities, and seeing the percentage of teachers who have experience in using computers for teaching, few teachers seem to have experience in evaluating ECPs. Thus, another important need was that teachers had to learn how to evaluate ECPs.

With the participation of 28 Grade 10 – 12 teachers and 219 learners, I set about developing a comprehensive evaluation scheme by evaluating a CAA programme that can address problems in assessment, and Zadarh that can address problems in interest and conceptual understanding, in 23 South African schools. The evaluation scheme attempted to be sensitive to learning, and to educational needs in disadvantaged communities such as those found in SA. It is important to note that teachers are likely to evaluate programmes against what they consider priorities. For example, a teacher who believes that a science laboratory is urgently needed might look at a computer or ECPs as a luxury.

The second level of this evaluation

The second level of this evaluation sought the values teachers attached to CAA and Zadarh. The evaluation process was also researched. Therefore, the rest of this thesis is about this second level.

THE STRUCTURE OF THIS THESIS

This thesis comprises seven parts. The previous section, the Introduction, was Part I. Part II follows and provides the theoretical considerations on the nature of science (NOS), biology, and learning strategies, which are predominant in South African schools.

Part III provides an outline on the possibilities of using Instructional design (ID) to solve problems in science classrooms.

Part IV indicates the process of evaluation as well as the problems that had to be solved in the evaluation process – i.e., besides evaluation of the programmes, the evaluation process itself was under scrutiny with regard to its validity and worth to participants.

Part V presents data and the analysis of data on the evaluations teachers and learners made of the two programmes. The discussions of these data appear in Part VI.

Part VII gives the conclusion and recommendations.

References, appendices, and manuals follow. Three manuals include an evaluation scheme; guidance on playing Zadarh, and; an outline for teachers on evaluation. The questions participants asked guided the contents of these manuals.

PART II

THEORETICAL ORIENTATION AND CLASSROOM PRACTICES IN SOUTH AFRICA

Understanding the philosophy of the subject, the learning theories, and how teachers apply these is important to evaluate whether a programme represents the subject accurately and how far the programme is applicable in class.

> ***Worthwhile programme evaluation must consider the underlying theory regarding the programme itself in order to connect meaningfully a programme's features and mechanisms with outcome measures***
> **(Chen as cited in Hickey & Zuicker, 2002: 541)**

SOME OF THE SCIENCE PHILOSOPHIES AND METHODS IN SOUTH AFRICAN SCHOOLS

Why worry about methods and philosophies in science classrooms?

Chen's quotation on the previous page informs the need for this section in this thesis. I believe that it is necessary to know the philosophy on which one bases science education because it clarifies the reasons why that person teaches science, and why that person chooses to use some methods to teach science. In the context of this research, the philosophy of science should be a foundation for designing Educational Computer programmes (ECPs). The evaluations include seeing whether, and how, science teachers and the two ECPs deal with inquiry, processes, and methods.

This section deliberates on some popular science philosophies operating in South African classrooms, as well as scientific inquiry, processes, and methods that are often emphasised in South African science curricula. I start by giving basic definitions and then describe how the South African science teachers practice these in classrooms, apparently unaware of the philosophies.

Philosophy of Science or Nature of Science?

Introduction

The Nature of Science (NOS) (i.e., the epistemology of science - the origins, nature, and content of science) is often used to describe the philosophy of science. This section explores the major theories, used in South African schools to obtain what is assumed scientific knowledge.

Formerly known as 'Natural Philosophy', 'science' comes from 'scientia' (Latin), meaning 'knowledge', that can be proven (Ross, 1998; Pecorino, 2001). Science was seen as heroic, truth seeking, autonomous, for a privileged few and above the general public. But a social agenda for science started to develop such that what constitutes science is a public debate, and is explained variously (E.g., McComas, Clough & Almazroa, 1998: 4), although that unchallengeable 'truth' image of science has persisted in some classrooms.

Major philosophies of science apparent in South African classrooms

Rationalism

Rationalism is reason, especially upon the Euclidian geometric system, deduced from abstract innate ideas or prior knowledge (*a priori*), independent of sensory experience (O'Hear, 1989; Popper, 1974; The American Heritage Dictionary of the English Language, 2000). Rational applications in science include the use of equations or models to predict phenomena.

Many theorists such as Dede, Salzman, Loftin, & Ash (1997), Stratford (1997: 4), Sanders (2002), and Sanders & Khanyane (2002) have noted that models form the basis of teaching science, and I think that models can translate into mental schema or abstractions, if learners have an understanding of the elements from which a model is constructed. This kind of modelling the world for learners and challenging them to extrapolate such models is arguably characteristic of rationalism.

One of the problems with rationalism is that results of an experiment might not fit prediction.

Empiricism

In contrast to rationalist innate knowledge, empiricism assumes that brains are blank (*tabula* rasa) until they are exposed to some experiences (*posteriori)* (Medawar, 1969: 27; The American Heritage Dictionary of the English Language, 2000). Empiricism recommends experiments and testing of hypotheses (Popper cited in Kuhn, 1974: 800), and extrapolation of data into laws or principles of science (O'Hear, 1989) such as Ohm's Law. Deriving, verifying, or proving empirical relationships in school experiments are assumed to be doing science.

The problem with empiricism is establishing the certainty of a phenomenon, or the accuracy of observations and measurements. Empiricists do not account for fallibility; that is, how senses lead us to error (Medawar, 1969: 31).

Positivism

A seeming extension to empiricism, positivism (or logical empiricism) relies on precise, certain and objective measurements, and rejects subjectivity and human ideology, history and intervention or intuition (Stockman, 1983: 30; Trochim, 2002). According to Comte (the founder of positivism), constructively valuable knowledge is based on facts, which arise from useful, certain, and precise data (Stockman, 1983: 30). Such preciseness is achievable if we assume that objects exist independent of any subject, such that the goal of science is simply to describe the objects through experience and or observation of phenomena (Trochim, 2002). That is, knowledge that cannot be observed is not scientific. Positivism is therefore hostile towards religion and metaphysics (because these are immeasurable). Hence, positivists emphasise the externality of the laws of nature to the observer: knowledge is out there and is perceived the same way by every careful observer (O'Hear, 1989: 14 - 19). We achieve objectivity by continued observations, bearing in mind that these may turn out wrong data (Scheffler, 1967: 1).

Objectivism in positivism

Objectivism assumes that there can be consensus on findings and meanings through observations. It relies on data, *which cannot be doubted* (Stockman, 1983: 30, 38), through triangulation, verification, control, impartiality, sampling a reasonable proportion of a population, and statistical wizardry (Scheffler, 1967: 2; O'Hear, 1989: 6). Objectivism relies on spatial conceptual frameworks (Schlick, 1925: 530 – 531). Thus, the essential properties of objects are knowable and relatively unchanging, and the world is real, and can be modelled for the learner (hence, the physical and mathematical models for science concepts).

Other positivisms practiced in classrooms

Positivisms are not clearly demarcated from each other. I deal with those that seem to influence South African science classrooms.

a. Logical positivism

Logical positivism is attractive in as far as it dilutes the importance of objectivism. Logical Positivism draws rationalism (logic, instinct, etc) and empiricism together (Rosenblueth, 1970: 4; Medawar, 1969: 28), and earns the claim that science is 'hypothetical-deductive' (Wellington, 1994: 24). Unlike radical positivism, logical positivists recognise the limitations of human beings to the extent that the scope and acquisition of knowledge is limited by scientists, rather than physical reality, and by previous knowledge and rules, which are used to interpret experiences (Cobern, 1996: 302-303). However, findings are never final and are always subject to question and doubt (Einstein, 1940: 253) such that one part of the theory is logical, while the other part is empirical (Spector, 1965: 44 – 45).

b. Post-positivism

There is an apparent similarity between Logical Positivism and Post-positivism – both uphold hypotheses and deductions as important parts of science. Except, Post-positivists believe that human fallibility can be minimised through repeated experiments (Trochim, 2002). In other words, objectivity is increased by triangulation (O'Hear, 1989: 61-62; Feigl, 1949: 11-12). According to Troachim, one of the common forms of post-positivism is *critical realism,* which takes observation as fallible and theory as revisable. A critical realist is *critical* of the means by which logical positivists obtain knowledge to the extent that scientists persistently try to get to the truth. Thus, critical realists try to be objective (Bodner, 1986: 874). Even though that goal can never be achieved (Trochim, 2002), the truth or real world exists regardless of our perceptions (Bodner, 1986: 874). Feigl likens the survival of scientific theories through triangulation to the survival of a species in evolution. Thus, objectivity is not the characteristic of an individual; it is inherently a social phenomenon or rather a social construction. Science advances by consensus and revision (Linnerman, Lynch, Kurup, & Bantwini, 2002: 205-210).

Counter arguments against 'normal' philosophies of science

Arguments against objectivism

Arguments against objectivism are abundant, and include that no number of experiments or verifications can establish positively a reality, and that multiple observations of the same phenomenon might differ (however slightly), such that some illusive and subjective nature of reality exists (Cobern, 1996: 302; Pajares, 1998; Geelan, 2000).

Some theorists believe that objectivism survives by 'scientific' models or spatial frameworks and statistics, by which we attempt to model nature (Schlick, 1925: 530 – 531; Scheffler, 1967: 2). Thus, modelling is an important strategy in science teaching. Penner (as cited in Jonassen, Howland, Moore & Marra, 2003: 190) mentions two basic forms of models: physical models, which are visible or concrete, and; conceptual models, which are not visible but thoughts. In this thesis, a model is considered to be a construct that imitates the real concept or *natural system* and its interactions (Stratford, 1997: 3-4).

Popper and daring hypotheses

Popper observed that a *verifiable statement* (or law) has to be written in absolute terms (Stockman, 1983: 24; Scheffler, 1967: 5). Popper starts his argument by pointing out that an absolutely and irrevocably true statement forbids particular occurrences (Popper, 1974: 962 – 963). I.e., it leaves no room for doubt.

Popper does not see any science in doing something that will surely happen (i.e., something that can be verified with certainty). Thus, one can argue that Popper does not support verification of or proving science 'facts' or 'truths' as presented in classrooms. Verifying truths requires one to follow prescribed methods without question since a modification of methods can lead to another truth. Prescribed methods are used in classrooms by means of worksheets.

Instead, Popper advises scientists to try to falsify those 'truths'. That is, Popper demarcates science knowledge from other forms of knowledge, by *falsifiability* of a hypothesis (Popper, 1974; O'Hear, 1989: 56; Pajares, 1998; Geelan, 2000). An attempt to falsify laws of science opens science to scrutiny and to new ways of doing science. One way of falsifying a theory is by stating hypotheses that are daring in light of what is considered to be the truth (Popper, 1974: 978-984). Popper points out that *great scientists stated refutable and falsifiable hypotheses* (and many were rejected: E.g., Charles Darwin). In other words, teachers should rather help learners to look for data that falsifies theories than look for data that proves the theories. Alternatively, teachers should relax the rules they give to learners during practical work so that learners come up with their own truths and ways of establishing truths. In this way, Popper's recommendations would fit constructivist approaches.

Kuhn and revolutions of paradigms

I think that Popper and Kuhn were in agreement in terms of criticising 'normal' science. Popper advised against clinging to truths or to verifying laws, which implied following prescribed methods, and Kuhn believed that these prescriptions presented in form of paradigms can never lead to new knowledge, unless they were violently changed.

However, Kuhn states that he disagreed with Popper in that Popper substituted verification with falsification (Kuhn, 1974: 799). Verification and falsification are complementary because, according to Kuhn (1974: 813), while verification confirms a theory, falsification serves to show the incompleteness of that theory. Hence, Kuhn instead suggests that advances in science happen from paradigm shifts, rather than falsifications.

Different authors define a paradigm variously:

> *... achievement that for a time provide model problems and solutions to a community of practitioners* (Shepere, 1984: 37); *... a collection of beliefs shared by scientists; or a set of agreements about how problems are to be solved* (Ross, 1999); *... a framework within which scientists do their day-to-day work* (O'Hear, 1989: 65).

All these definitions imply that scientists follow certain agreed-upon rules. That is, paradigms arise because scientists form their own social system (Gardner, 1975: xiv), which is governed by a set of rules (O'Hear, 1989: 65). Thus, objectivity is not the characteristic of an individual; it is inherently a social phenomenon or rather a social construction. Normal science *advances by consensus and revision* (Linnerman *et al.*, 2002: 205-210). The scientific community rejects unconventional methods of obtaining knowledge, and insists that findings build upon established truths. Note for example, how learners who fail to get the correct data are punished with failure and are eventually expelled from science classrooms, instead of finding out how they claim to know what they write on paper.

However, extreme adherence to paradigms restricts falsification, and pollutes minds and senses with pre-conceptions because paradigms determine acceptable scientific techniques, data analysis, and explanations, and therefore determine data itself (Medawar, 1969: 29). As with verifiable statements, paradigms tend to suppress fundamental novelties (Shepere, 1984: 38), because paradigms condition the theoretical world-view, and observations (O'Hear, 1969, 69). Medawar (1969: 25) concludes that *innocent unbiased observation is a myth,* and that the intellectual processes during investigations are themselves the grounds that justify knowledge. Suppression creates false peace and order in the scientific community and no new developments happen during such peaceful rule-bound periods (Shepere, 1984: 37).

For new ground-breaking knowledge, the peaceful interludes are *punctuated intellectually by scientific wars or revolutions* (Ross, 1977). The revolutions challenge the orderliness imposed by paradigms, especially those that embrace objectivism and verification (Scheffler, 1967: 3). Kuhn described these wars as "*the shattering complements to the tradition-bound activity of normal science*" (Pajares, 1998). Thus, Kuhnian science does not develop cumulatively because revolutions replace one conceptual worldview by another (Geelan, 2000).

Intuition

It was mentioned in the above that positivism does not accept knowledge that cannot be proven by scientific methods. One of the dimensions of science that is neglected as a result of that is intuition.

Popper and Kuhn's arguments arise from the claim that scientists tend to adhere to truths and paradigms. However, Ross (1977) observes that scientists obtain knowledge without strict adherents. This is possible because the boundaries of paradigms are hazy (Shepere, 1984: 40). Among the processes that do not fit in any paradigm is intuition. Creativity is often intuitive and subjective, but science classrooms and assessment do not encourage it.

Intuition, together with adventure and imagination, is responsible for the development of ideas, new cognitive structures and concepts (Ross, 1977; Adey, 1987: 19; Wellington, 1994: 24), and accounts for achievements that cannot be explained logically or precisely (Ross, 1977; O'Hear, 1989: 10). Furthermore, in contrast to O'Hear's (1989: 10) criticism of Aristotle's notion of intuition, Medawar (1969: 46 - 55) points out that *imaginative or inspirational processes enter into all scientific reasoning at every level,* such that inventors acknowledge the role of these processes. The development of the Watson-Crick model of nucleic acids is a good example of intuition (Wade, 2003). However, we teach the Watson-Crick model as if Watson & Crick developed it from logic and prescribed methods alone. Problems can arise when such intuitively developed concepts are taught purely objectively and logically.

Medawar (1969: 55-57) mentions four forms of intuition that can have implications in analysing a computer game such as Zadarh:

- *Deductive intuition: perceiving logical implications instantly, and seeing what follows from holding certain views*

- *Inductive intuition: creativity or discovery, which is thinking up or hitting on a hypothesis*
- *Instant apprehension of analogy: real or apparent structural similarity between two or more schemes of ideas (wit)*
- *Experimental flair or insight*

The affective domain in science
Laderman (1998) suggests that attaining an understanding of scientific inquiry, and the NOS should include the "affective" aspect of learning. Theorists (E.g., Gagné, 1985; Trollip, no date; Tamir, 1996: 109; Harlen, 2000), show that the affective domain to involve values and feelings such as the willingness to collect and use the evidence (respect for evidence), willingness to change ideas in the light of evidence (flexibility), and willingness to review procedures critically (critical reflection).

Scientific inquiry, processes, and methods

It is possible that the confusion concerning what constitutes science has contributed to problems in science education, given the varied definitions of science with no agreement between them. Although theorists agree that science inquiry represents what scientists do, they disagree on the processes and methods. In traditional schools, scientific inquiry therefore appears to be a mixture of rationalist, empiricist, and positivist approaches.

Scientific inquiry seeks scientific knowledge (Aristotle quoted in Wartofsky, 1968: 291), and inquiry explains science (Savage, 1998: 51; Thomson & Stewart, 2003: 161). Several theorists (Thomson & Stewart, 2003: 161-162; Lederman, 1998; Savage, 1998: 51; Wartofsky, 1968: 205; Munro, 1975: 220) state that the activities during scientific inquiry are science processes. *When cognitive development is operationalised in terms of Piagetian reasoning patterns and related tasks, considerable overall with science process skills is evident* (Brotherton & Preece, 1996: 65). These include designing questions, conducting investigations, critically analysing the claims and assumptions, making predictions, understanding limitations and the implications of the investigation and data, problem-solving, discovering properties of things and relations, etc.

Science education emphasises learning as a process (Tobin, Tippins, & Gallard, 1994: 46; Gibbons, 2000: 3), and several countries have placed a heavy emphasis on process-oriented hands-on learning in science education (Kumar, 1994: 59). However, there is no clear definition of what constitutes a scientific process to the extent that other disciplines, such as information technology, can lay claim to some of them. More importantly, processes cannot be demarcated from each other as many of them can happen concurrently (Harlen, 1993: 28).

A scientific method is a sequence of processes (Lederman, 1998; O'Hear, 1989: 12). That is, methods are seen as ways of ordering and testing knowledge in the course of scientific inquiry (Henry 1975: 62; Wartfosky, 1968: 205; Linnerman *et al.*, 2002: 205-210). Depending upon the sequence, methods can be hypothetical –deductive: that is, we can start by stating a hypothesis, carry out a study, and then draw a conclusion. Or methods can be inductive: that is, we can start from simple evidence to design a study and derive generalisations. For example, see Spector (1965: 44), Medawar (1969: 23-26), O'Hear (1989: 12), and (Kneale, 1949: 353).

Dewey (as cited in McComas, Clough, & Almazroa, 1998: 7) advised that understanding science methods was better than acquiring science knowledge. However, the popularity of methods could have drawn from a myth, especially in political circles, that there is a unique 'scientific method', based upon *classical experimental design* (Lederman, 1998), characteristic of empiricism, positivism, and causality (Russell, 1929: 387; Medawar, 1969: 14-21; Pecorino, 2001), with an assumption that the path from evidence to theory is logical if one uses prescribed methods (Kneale, 1949: 353; Sutton, 1996: 2; Einstein, 1940: 253). Science students have been faced with proving empirical formulae, for example, in chemistry. Formulae in relate variables or show how a change in one variable causes another variable.

Russell (1929), Feigl (1953), Nagel (1951; 1971), Wartofsky (1968), Medawar (1969), among others, give extensive accounts on causality. Causation (as understood from Galileo, Hume, and Mill) is *defined in terms of predictability according to a set of laws* (Feigl, 1953: 408; Russell, 1929: 390), for example, using equations and statistical analyses, by which it is assumed that similar conditions achieve similar results, which can be extrapolated. However, it is difficult to get real life experiences for all learners. Therefore, Bacon suggested invented happenings, or contrived experiences called experiments because nature is so vast (Medawar, 1969: 35). The purpose of experiments is discovery and testing, which is done by systematically relating instances, and formulating expectations (Wartofsky, 1968: 206). Experiments can follow hypothetical, inductive, or deductive methods (Medawar, 1969: 34 – 39). At school, experiments isolate dependent and independent variables, and aim at deriving their relationships.

Henry (1975: 61 – 74) reports that teachers from different parts of the world have come to expect many outcomes from experiments. Teachers hope that their learners would understand the nature of science by practicing science processes such as testing hypotheses, problem solving, and observing, through doing experiments in laboratories. Thus, p*ractical work is now a standard feature of secondary school science, with the belief that science belongs in the laboratory* (Jenkins, 1999: 21-22). It is an unfortunate belief that science belongs to the laboratory.

Science and culture

Science education faces problems regarding the tensions between science and culture. These tensions are expected if one believes Ormell (1980: 87) who states that any subject can act as a carrier wave for value systems.

Firstly, if we take culture as the norms, values, beliefs of a group, science qualifies to be a culture that scientists as a group practice, independent of other cultures. However, if science is a human attempt to understand nature, then every culture has its science such that scientific knowledge is socially negotiated (Cobern, 1994: 63; Cobern, 1996: 299-300; Jegede, 1998: 156). *Science is not a culturally independent phenomenon, it comes with a way of viewing the world and with certain values attributed to the kind of knowledge it deals with* (Kuiper, 1998: 1). Wood-Robinson *et al.*, (1997: 182) seem to take this position in the statement ... *individuals need to have an understanding of science as a cultural achievement of modern society.* Science then would antagonise cultures, which are still to be modern (or even those that are post-modern).

Thus, 'modern society' or the 'western world' is a culture whose philosophies I have debated earlier. Such science might be problematic in non-Western cultures because its symbols are contextually and culturally embedded (Lincoln & Denzin, 1994: 579;

Moje, 1995; Atwater, 1996; Ogunniyi, 1996, 1997, 1984; Driver, Asoko, Leach, Mortimer, & Scott, 1994: 5-7; Lederman, 1998; Cobern, 1996: 295-296; Makhurane & Kahn, 1998: 28; Henderson & Wellington, 1998). For example, learners might lack the linguistic competence to express their understanding in scientific terms.

Contrary to that view, philosophies such as empiricism and positivism, through the science method tend to make science an enterprise that is independent of context and culture. According to Cobern (1994: 65) as well as Makhurane & Kahn (1998), such science transcends all cultures, whereby the results of scientific inquiry are the same regardless of context and culture. This belief seems to be popular, among teachers (Atwater, 1996: 828).

Although urbanisation and globalisation have removed some barriers between nations and cultures, it appears that culture and surroundings (E.g., social and economic factors) extensively influence an individual's conception of 'science'. Teachers as well as instructional designers should note that, awareness of the learners' cultures (including language), and how these relate to science can improve understanding science (Cobern, 1996: 295-296; Rollnick, 1998: 87), and also that this approach does not easily fit into established science paradigms. Hence, Lewin (reported in Cobern, 1996: 29) notes that there is still a long way to go in developing ways of representing science that are not foreign and expert, and which are culturally unsympathetic.

Science in South African classrooms

In this section I attempt to draw conclusions on what kind of science philosophies operate in South Africa on the basis of my experiences and observations of science classrooms (for example, classroom observations in Part I of this thesis).

The school science curriculum in SA recommends inquiry, processes and methods, but no science philosophy is recommended or mentioned. Nonetheless, even though it has failed in certain cases, this approach, in which scientific knowledge is based upon observations, is characteristic of empiricism. One of the science teacher's struggles is helping learners establish with certainty and accurately a phenomenon through observations and measurements. *For example,* experiments on transpiration require learners to accurately record the rate of water loss under different conditions.

Learners are then expected to prove laws that govern plant-water relations. However, teachers run out of ideas when learners record different observations and measurements, since empiricists do not account for fallibility (Medawar, 1969: 31).

The South African science syllabus includes skills of reading instruments accurately, and of presenting data adequately. Where learners carry out project work, teachers advise them to obtain data using a variety of methods – this is triangulation. These measures increase statistical validity of data and objectivity (Scheffler, 1967: 1-2; Stockman, 1983: 30, 38; O'Hear, 1989: 6). It is characteristic of logical empiricism (positivism).

Few teachers do practical work

Despite recommendations for scientific inquiry in curricula, many teachers believe that it is their duty, and that they are able, to model reality for the learner and supply objective truths to the learner, often using 'facts' they find in the textbooks (see preliminary research, Part I). In other cases, equations in textbooks are popularly used to extrapolate deductions into predictions. The atomic structure as taught at Grade 10, and the structure of the DNA in Grade 11, are examples of models that are logical and abstract deductions.

There are many reasons to account for the non-practical modelling approach. These include that, expectedly, there are cases when experimental data does not fit textbook predictions or statements, partly because laboratory equipment in schools is rarely able to achieve the accuracy required even in the hands of experienced teachers. Even more unfortunate is that some science teachers in SA have been trained to believe in textbook facts, even where data does not support those 'facts', and so would rather exclude such experiments from instruction. The other problem is either lack of equipment to do practical work or the teachers' problems in practical skills and conceptual understanding.

The nature of the South African examination that excludes the assessment of practical skills has also helped in encouraging teachers to chalk and talk in class (see preliminary research, Part I).

The result of all these problems is that the teachers venture into logical explanations and abstract deductions that are characteristic of rationalism. This is not to say that rationalism is useless, but that it is misused in South African classrooms.

Dissatisfaction with positivism or objectivism

There are arguments within SA emerging from dissatisfaction with objectivism, which augurs well with the liberation struggles against apartheid and its education. C2005 is in line with the critical theory or critical realism, which, according to Nichols & Allen-Brown (2001) and to Bodner (1986: 874) challenges the taken for granted and control of subjects, and according to Giroux (as cited in Atwater, 1996: 823) challenge *hegemonic ideologies*, and gives a voice to subordinate groups, who in the case of science education in SA continue to be the Black people. But it is arguably also dissatisfaction with behavioural strategies – I will discuss behaviourism later. C2005, like critical theorists, discourage the notion of precise or absolute truths, as well as objectivism, and the systems that claim to bring these about because it acknowledges realities in different cultures and contexts.

Challenges to approaches teachers use in SA to teach science

The overriding point is that teachers follow verbatim curricula and textbooks that the DoE provides. That is, curricula have become rules to be followed.

In 'normal' science classrooms, the DoE advises that learners become experts in the science processes, methods and processes, so that they can find employment. School products are rarely expected to create jobs. It should be appreciated that one legacy of apartheid education was to create labour reserves in homelands and townships (Christie & Collins, 1984). The kind of labour required was that, which could follow instructions verbatim, without question or innovation. Thus, the job market in SA prescribes procedures, especially those that verify some selected 'relevant' laws, and 'accepted' explanations for phenomena. (This influence can also be seen in light of objectivity and behaviourism). Similarly, Atwater (1996: 830) notes that *power and control are two important elements in determining the science curriculum*. This conformist or controlled situation or rules in curricula and science classrooms that define objectivity makes science a closed system, which, in my opinion, Popper and Kuhn have questioned, and which challenge teachers to change towards more open constructivist science teaching. These open environments allow exchange of ideas and knowledge between teacher and learner, and offers opportunities to the learners not only to state daring hypotheses (hypotheses from the point of their world view or culture), but also to use unconventional methods, including those embedded in their cultures, to test scientific theories or to test their worldviews.

Science and culture in South African classrooms

Manzini (2000: 21) who was a science teacher in South Africa then sums up his experiences of teachers' problems in relation to culture:

> *...teachers seem oblivious of the cultural bias of the present curriculum. They do not think critically about the concepts, aims, approaches, and resources it advocates. They merely try to transmit the curriculum voetstoets, thereby perpetuating the status quo of perceived abysmal performances among African learners in science. They find themselves as accomplices in the cultural genocide, albeit inadvertently.*

With that kind of feeling, it would be expected that culture had to become a topical issue in the South African science curriculum – C2005, and rates highly in the critical outcomes. There is a feeling that science concepts might be better understood in the context of a learner's culture.

Popper and Kuhn's arguments apply to the clashes between paradigms in cultures against paradigms in science. Unlike much of the Western world, SA is a combination of deep-rooted cultures with very fundamental differences. Language differences are examples, where translation from one language to another, or translation of science concepts to many of the languages is difficult, and inevitably challenge the preciseness of concepts and paradigms.

Other culture-science conflicts relate to phenomena such as AIDS and lightning. For example, some tribes have beliefs on lightning to the extent that I have had interesting debates, contradicting electrostatics theory, when teaching how lightning occurs. My experience is consistent with the assertion that culturally neutral observations are rare (Atwater, 1996), but is contradictory to the assumption that a scientist should preserve a complete freedom of mind (Bernard as cited in Duhem, 1953: 235). Either observation is selective on the basis of held theories and methods or one looks at science from a background of his/her culture (Lederman, 1998).

Problems from the Nature of Science (NOS), methods and processes in class

From the above, it is apparent that there are problems with the NOS in science classrooms in SA. This can be expected, since none of the 104 teachers in a Carnegie CASME workshop during 2004 (page 20) was able to discuss the NOS. Teachers claimed that they had not studied NOS formally. This is in agreement with other findings (E.g., Mkhwanazi, Mkhwanazi, Rollnick, & Bradley 2002: 260-264; Fabiano, 1998: 137; White, 1996: 761).

Without knowledge about the NOS, teachers do certain things without understanding why. Hence, science is distorted in class (Tinker & Papert, 1988: 1-2; Gilbert & Watts, 1983; Sutton, 1996:1; Harlen, 1993: 2; Makhurane & Kahn, 1998: 23), and science education has mishandled NOS as well as scientific inquiry (Laderman, 1998).

The disagreements on which science philosophy is most appropriate (Popper, 1974: 1015; Medawar, 1969: 24), brought about by the multiple definitions of the NOS (Brodbeck, 1953: 3) confuse curriculum designers and teachers on the choices of what is to be learnt, and how it ought to be learnt. Obviously, an agreement on the meaning and choice of one philosophy would domesticate science and restrict the richness of ideas on science (as Kuhn feared), loosing out on the much desired diverse viewpoints of science that match the complexity of nature, which science attempts to explain. However, domestication of certain philosophies already exists in South African classrooms, although teachers might not be aware of it.

Conclusion on philosophies of science in South African classrooms

Rosenblueth's (1970: 2) view that there seems to be no agreed-upon definition of science makes sense. Nor is there a preferred philosophy. The philosophies that I have outlined seem to work in South African schools concurrently, albeit without awareness or conscious intention on the part of the teachers. Additionally, I agree with Mgujulwa & Kenyon (1994: 260), that teachers simply teach the way they were taught.

This situation poses a challenge to instructional designers in that they might not subscribe to a science philosophy and yet have to convince end users that a programme is a science programme. In particular, the challenge is choosing modes of inquiry, or methods (sequence of processes) that are acceptable to the DoE and the body of scientists, understood by science teachers, and compatible with the way teachers conduct their lessons. One way of reducing this challenge is to get the DoE and teachers involved in the design process, for example by getting them to evaluate the programme formatively.

The situation also leads to wonder whether it is necessary to stick to or mention a philosophy we use. Sticking to a philosophy might domesticate science, while abandoning philosophies leaves teachers of science defenceless against the need to justify the methods they use for learners to know science. The other problem is determining how these philosophies antagonise or promote our cultures. Finally, we need to know how each philosophy relates to the different learning theories.

These arguments have existed for centuries, and will most likely continue indefinitely. However, science teachers and instructional designers should think about the impressions they create in the minds of their learners about the NOS.

Concluding Popper and Kuhn's challenges, I suggest that some computer programmes can be used to encourage constructivist strategies by which learners can state daring hypotheses and in which rules or paradigms as set in science classrooms can be challenged.

TEACHING STRATEGIES IN SOUTH AFRICAN GRADE 10-12 SCIENCE CLASSES

Why include learning strategies?

I outlined the philosophies of science that seem to be operating in South African classrooms in the previous section because I believe a valid evaluation should consider the theoretical grounding of an instructional design. Similarly, I believe that a valid evaluation of an ECP requires the knowledge of the underlying epistemological assumptions of the programme, and how these relate to the teachers' practices. Teachers and instructional designers plan lessons upon their belief on how learning happens. In this section, I examine popular teacher's practices and how they relate to some aspects of learning theories.

Again, these are based upon my experiences as a teacher in SA, preliminary studies, and class observations in the course of my work, and are supported by other researches in literature.

Learning theories

This thesis considers three major learning theories: behaviourism, cognitivism, and constructivism, aware that these have been redefined in so many different forms or branches. I give the foundations of these three learning theories below.

Behaviourism

Behaviourism is learning by associating a desired behaviour with extrinsic motivation or environments that are provided through reinforcements or rewards (Fosnot, 1996: 8), and concentrates upon observable indications of learning (or behavioural change) achieved through combining a sequence of stimulus and response 'cause and effect' relationships from simpler to complex behaviour (Conway, 1997: 1 – 2; Child, 1997: 10). Learners are assumed blank on a concept that they have not experienced before (endorsing empiricism), and so the teacher or instructor designs a learning process or a programme that 'fills-up' learners with knowledge (Winn, 1997; Child, 1997: 10). Furthermore, reward should follow quickly when the correct response appears. This feedback on progress is motivating and provides to the learner an opportunity to discover stimulus discriminations for the most likely path to success. It is notable that positivism also endorses the principle of causality except that it does not directly acknowledge it in human behaviour.

A behavioural science classroom starts from simpler to more sophisticated processes, and rewards the ability to follow practical procedures. However, White & Tisher (1986: 876) conclude that there is no agreement on the influence of behaviourism upon science learning.

Cognitivism

a. Introduction

Craig, Mehrens & Clarizio (1975: 131) give some of the important differences between behaviourism and cognitivism. For example, cognitivism:

- *Views learning as a change in knowledge, not a change in response*
- *Recognises mental processes such as purpose, insight and understanding*

- *Takes learning to be more related to understanding of relationships in the present situation than to past stimuli*

Unlike behaviourism, cognitivism explains mental processes, and intrinsic motivation (i.e., what makes us do things without external persuasion).

b. Experiential learning

The cognitive explanation of experiential learning seems to be a direct off-shoot of John Dewey's philosophy of education, which recommends an education that is useful to an individual's life. Kraft & Sakofs (1988) explain that experiential education involves actively engaging people in experiences that will have useful consequences.

Experiential learning is durable, transferable, self-regulated, and applicable in solving problems in different contexts (Lawton & Hooper, reported in Mwamwenda, 1993: 71; Wollman, 1990: 555). This stance assumes that learning is a discovery and problem-solving activity. Adey (1987: 17-19) explains how experiential learning could be achieved. Adey advises that learners need experiences that can cause the development of logical thinking, and that this is characterised by the 'schema of formal operational thinking', and the ability to handle fluently: *Control and exclusion of variables; Compensation and equilibrium; Combinations; Frames of reference; Ratio and proportion; Correlation; Probability; and Conservation involving models.*

These variables are abundant in play, according to Gredler (2001: 521), and practical work, because these provide for discoveries and experiments with relevant knowledge, instead of hearing or reading about the experiences of others. One easily reflects upon one's own experiences.

c. Situated cognition and learning environments

Obviously, every environment imparts or requires some knowledge and skills. Situated cognition concerns the relevance and transfer of information learnt in a specific situation to different situations in life (Anderson, Reder, & Simon, 1996). Knowledge or skills learnt in different situations can also be applied in a specific situation. Either way, the environment in which learning happens is important. Such an environment must provide the necessary tools, resources, and the freedom to a learner to explore and use these for specific knowledge (Jonassen, 1991: 11-12).

Situation-specific learning enhances self-concept and a feeling of control over one's own success (Weiner cited in Rieber, 1992: 99).

d. Cognitive load theory

Cognitive load is the capacity of the brain in handling processes and knowledge (Wilson, 1995b), and is influenced by the amount and type of processing required (Hannafin & Rieber, 1989: 96). Simultaneous multiple elements compete for mental processing and impose a heavy cognitive load, which then threaten successful learning, although people can adjust their speed of processing to the amount of instruction provided (Hannafin & Sullivan, 1995).

Hannafin & Rieber (1989: 96) explain that the meaningfulness and familiarity of the lesson content as well as of the information codes are among the factors, which reduce cognitive load. These factors are directly informed by prior knowledge such that lessons ought to start with a learner's knowledge. Therefore, a teacher or instructional designer has to establish the rate at which a learner adequately works through experiences and tasks because the rate at which schema are formed or modified are limited (Wilson, 1995b). That is, an optimum rate of providing information has to be worked out, which does not crowd the learner. In this regard, Sweller (as cited in Wilson, 1995b) notes that it is important to analyse the number of elements requiring attention, and therefore to use single, coherent representations that learners can focus attention to rather than split attention between two places: E.g., between a diagram and the text. But at the same time a teacher has to eliminate redundancy between representations, for example, by providing opportunities for problem exploration. Exploration is more interesting with multimedia such as animation and audio narration. Besides normal life experiences, computers can provide these environments for exploration.

e. Cognitive conflict and conceptual change

Another important aspect of cognition is that it explains how learners transform their thinking or constructs. Practical experiences are supposed to transform the way a learner thinks. Tobin & Jakubowski's (as cited in Etchberger & Shaw, 1992: 412), Posner, Strike, Hewson, and Gertzog (1982), West and Pines (as cited in Wollman, 1990: 555), as well as Prawat (1992: page 4, para 3) explain conceptual change. Learners become intellectually disturbed through exposure to experiences (for example, through practical work and real problems), which challenge their knowledge. Learners will experience cognitive conflicts and realise that they need to change their constructs if they fail to solve problems (Wollman, 1990: 555). The teacher often has to support learners if they are failing to find alternative conceptions or strategies. One way of motivating learners is to help them diagnose their conceptual problems (Posner *et al.* as cited in Geelan, 2000: 4) or to assist them towards metacognition. Posner *et al.* (as cited in Wollman, 1990: 555; Geelan, 2000: 4) advise teachers to make sure that *the new conception is compatible with experiences and plausible in terms of* solving the problems generated by its predecessors, and that the new conception links old with other new conceptions. It is preferable that a conception is translatable between different representations: E.g., between verbal, mathematical, concrete-practical, pictorial, etc.

f. Multiple intelligences

Intelligence describes the way we perceive or comprehend and process information, and act on it. The traditional view is that intelligence is inborn and cognitive. Intelligence Quotient, or IQ test (now known as the Stanford-Binet Test) measures such intelligence, and institutions use it to place learners in particular academic streams or into particular careers. The problem with IQ is that it does not consider background, culture, and environment of a person, such that IQ is a measure of a person's ability at a given set of items.

The theory of multiple intelligences (MI) differs from the traditional IQ in that no two people have exactly the same profile of intelligences because each individual has a different genotype and environments. Various sources list a different number, and explanations of these intelligences while McKenzie (2001) speculates that there may be many more yet to be identified. Gardner (cited in McKenzie, 2001) lists many categories, some of which are currently recognised in the South African science curriculum:

- Visual/Spatial intelligence is the ability to learn visually and to organise things spatially. Visual aids such as charts, graphs, maps, tables, illustrations, art, and puzzles are useful.
- Verbal/Linguistic intelligence is the ability to learn and use language and arts through speech, writing, and reading.
- Mathematical/Logical intelligence is the demonstration of an aptitude for numbers, reasoning and problem solving.

- Bodily/Kinaesthetic intelligence is the ability in body movement and is demonstrated through games and hands-on tasks
- Musical/Rhythmic intelligence is the ability to understand sounds, and is demonstrated through songs, sound patterns, rhythms, instruments, and musical expression.
- Intra-personal intelligence is the ability for one to understand own feelings, values, and ideas.
- Interpersonal intelligence is the ability to cooperate with others.
- Naturalist intelligence is the ability to understand and deal with nature.
- Existentialist intelligence is the ability to understand philosophy related with the existence of humankind.

MI informs ID to recognise the possibility that different learners might prefer different modes of learning. These are similar to Gagne's model of Information Processing (Gagne, 1985). One of these, spatial cognition, is particularly useful in virtual environments.

g. Spatial processing skills and cognition

Spatial processing skills or cognition is also called spatial intelligence, and is one of the multiple intelligences. Osberg (1997) states that spatial cognition is an important building block to general cognition because cognition is predicated on the interaction of individuals' sensory-motor and neurological systems. These interactions, in order of increasing complexity, include, tactility, vestibular functions, kinesthesia, proprioception, ocular motility – visual/motor integration, laterality, binocularity – auditory/linguistic integration, and visual/spatial integration – auditory/visual integration. Osberg claims that these stages would begin to appear during the concrete stage (ages 7 – 11), and would continue to develop through and beyond the formal operations stage (ages 12 –16). This implies that learners in the South African Grade 7 and above would have fully developed visual/spatial abilities.

Spatial processing skills are an important component in cognitive development because they allow one to create meaning from manipulating objects. The improvement in spatial processing skills is related with Piaget's stage theory because it involves the comprehension of perspective, transformations, ordinal relations, classifications, probability, etc, all of which are higher order thinking skills (Cobert, 1996; Patterson & Milakofsky as cited in Osber, 1997). Fortunately, studies show a preference for visual rather than verbal learning styles (Osberg, 1997). It would thus be easier to increase cognition as well as problem solving by using visual and therefore spatial skills, which emphasise spatial relations, sequencing, classification, transformation and rotation, and whole-to-part relationships.

h. Weaknesses of cognitivism

It is argued that Piaget underestimated children's abilities and overestimated the formal thinking capabilities of adolescents, and that learners in different environments or cultures may not develop intellectually at the same rate (Biehler & Snowman, 1991: 70). Mwamwenda (1993: 71) reports studies in the USA where learners were at the concrete operational level that were supposed to be at the formal operational level. Psychologists are also worried about the abstractness and non-linear nature of cognition (Mwamwenda (1993). These weaknesses imply that cognition is difficult to access and measure, and introduce difficulties in evaluating a learning experience cognitively.

Therefore, Posner *et al.* (as cited in Wollman, 1990: 555; Geelan, 2000: 4) see a need for developing evaluation techniques for tracking the process of conceptual change in learners. Learners can evaluate themselves through metacognition; which, according to Hannafin & Rieber (1989: 96), is the continuous awareness of one's own cognitive processes, and the ability to select, and revise cognitive processing strategies.

Constructivism

Many theorists acknowledge the importance and prominence of constructivism in contemporary approaches to science learning (E.g., Bodner, 1986; Campbell, 1998; Tsai, 2000;). However, there are different interpretations and emphases of constructivism (Winn, 1997; Phillips as cited in Campbell, 1998: para 5; Yore, 2001; Fosnot, 1996). This claim is perceptible by the large number of *constructivisms* such as developmental, radical, trivial, cognitive, social, physical, etc. Hence, Cobern (1996: 301) notes that constructivism has created considerable confusion and controversy. These kinds might be complementary and/or applicable in different contexts. Further uncertainty about the meaning of 'constructivism' could stem from its dual application in psychology and in philosophy (Tsai, 2000: 193; Bodner, 1986: 2). Constructivism parades as a learning theory, explaining in more detail the cognitive conceptualisation processes, as well as a philosophy in education (Bodner, 1986; Conway, 1997: 2).

While cognitivism explains developmental stages and formation of concepts, constructivism focuses, in more detail, on the stages of schema construction and alteration (Cobern, 1996: 301; Tamir, 1996: 95), which result from interpreting experiences, or from solving problems (Biehler & Snowman, 1991: 429; Salviati quoted in Cunningham, 1991: 130; Driver *et al.*, 1994; Birenbaum, 1996: 6; Cunningham, 1991: 14).

Jegede (1998: 160) recommends constructivism in science because it is idiosyncratic or interpretative, informed by culture and social interactions (I.e., constructivism relies on prior knowledge [Yore, 2001]), and takes alternative concepts seriously (Kuiper, 1996). These alternative constructions could emanate from differences in experiences, for example, embedded in culture. Enabling construction implies facilitating and negotiation rather than imposition. Therefore, a simple transfer of foreign cultures might cause problems (Cobern: 1996: 303). After all, scientists interpret experience in light of personal knowledge, similar to the way one uses personal, culturally embedded knowledge to learn from new experiences. Similarly, Driver *et al.* (1994) and others note the importance of language used to label and validate constructs. We can take language as providing the codes with which to communicate thoughts and observations, as well as the interpretations or constructs. This autonomous nature of constructivist learners apparently links with its philosophical posture, since, as Splitter (1991:92) argues, *no mental construct can be isolated from the mental process which formed it.*

Constructivism should not be difficult to apply for a cognitivist (including those who have been used to using Bloom's objectives) because there are clear similarities between cognitivism and constructivism to the extent that Piaget, one of the most prominent cognitivists, endorses constructivism frequently (E.g., in Fosnot, 1996). One example of a similarity is the constructivist recommendation for practical work and experiential learning:

> *... one could develop a method of participatory education by giving a child the apparatus to do experiments and to discover things by himself ... '* (Piaget in Tinker & Papert, 1988: 3).

More fundamental to a constructivist way of learning science is the argument that *knowledge neither resides in the mind of the* knower nor in the environment being explored (Hickey & Zuiker, 2002: 540). At this individual level, where supposedly each learner takes responsibility for his/her learning and collaborates with others, *knowledge is 'stretched across' the social and physical contexts* (Cole, and Pea as cited in Hickey & Zuiker, 2002: 540). Activities between the 'social' and the 'physical' describe scientific inquiry.

'Conception' should be the focus of acquiring knowledge, and as with MI, this might happen differently for each individual (Kuiper, 1994: 280; Shymansky, Yore, Treagust, Thiele, Harrison, Waldrip, Stocklmayer, & Venville, 1997), and might lead to varied realities (Bodner, 1986: 874). At an individual' level, the terms 'learners' understanding' or 'alternative conception' are preferred in place of misconception (Kuiper, 1994: 280; Ogunniyi, 1997: 53); this legitimises a learner's interpretation of an experience or understanding of a concept, but could be problematic in validating a learner's errors. Yore (2001) suggests *inter alia* attention to individual learners, encouraging debate and dialogue, and continuous assessment in a constructivist class. The individual's uniqueness in thinking and meta-cognition: that is the independent, autonomous, and self-regulated individual capable of communicating and cooperating with others (Birenbaum, 1996: 4; Yumuk, 2002: 142), in my opinion, construes constructivism as interpretative. This autonomous nature of constructivism apparently links with its philosophical posture, since, as Splitter (1991:92) argues, *no mental construct can be isolated from the mental process which formed it.*

In Splitter's view then, Cobern's (1996: 301) argument that constructivism is a thought rather than a formal logical operation, is challenged. Rather, constructivism traverses intellectual operations (psychology) as well as constructs or thoughts (philosophy). As a philosophy, constructivism is a view on how we come to understand or know (Savery & Duffy, 1995: 31). Therefore, constructivism embraces the various means by which we seek knowledge. The question is whether it embraces the various means of scientific acquisition of knowledge, such as rationalism, empiricism, and positivism since these are types of scientific construction of knowledge, which individuals and certain groups of philosophers use.

However, it seems that constructivism can be described as a subjective form of *relativism* that provides *alternative epistemological bases* (Osberg, Winn, Rose, Hollander, Hoffman, & Char, 1997), and *focuses on the nature, methods, and limitations of human knowledge* (Atwater, 1996: 827). The subjectivity emphasises interpretativism. Hence, Osberg *et al.* argue that each of us could have a different way of explaining the same reality. These individuals' realities can be interactive and mutually constructive (Yore, 2001), to a socially agreed meaning.

Learner 'autonomy' requires development of capacity, which society or peer groups can provide variously (Vygotsky, 1962; Bodner, 1986; Jonassen *et al.*, 2003: 3), hence social constructivism. The social aspect of constructivism has regard for an active involved society in which the culture gives the learner the cognitive tools needed for development (Vygotsky, 1978). Conway (1997) explains that tools include *technology, which learners can use* to produce a product that they can share with other people, and that the type and quality of those tools determine the pattern and rate of development more than do Piaget's cognitive development theories. That is, apart from each of us constructing the world within the self (cognitive constructivism), we seek consensus through acknowledging that society collectively has to agree (or disagree) on meaning (social constructivism).

Learning strategies in the context of science education in SA

There is no officially recommended learning theory in SA, and, consistent with Ramsey's (1975: 96) caution that a teacher's knowledge *about teaching does not flow automatically from theories of learning*, it is difficult to identify the learning theory applied in a typical South African classroom. It is safe to say that there are instances of each one of the traditional learning theories in a single lesson, even when such a lesson includes assessment and play.

In the past, South Africa officially endorsed the philosophy of Christian National Education (CNE) and Bantu Education (BE). In line with fundamentalism, White South Africans used Christianity to argue that they alone can guide Black South Africans in matters of faith and education. The CNE argument augured well with Fundamental Pedagogics in the sense that learners had to be guided to true and useful knowledge (Enslin, 1984). Learners had to obey all rules to get this true and useful knowledge.

It appears that behaviourism was compatible with Fundamental Pedagogics, and its assumption that learners are blank slates (*tabula rasa*), in relation to new knowledge, and have to be led by well-sequenced stimuli and reward. (it is also noteworthy that *tabula rasa* and empiricism, share the assumption of a blank mind, until some experiences invade that mind).

In today's South African science classrooms, teachers use behaviourist strategies, sometimes inadvertently. Firstly, teachers 'teach' experiments (i.e., teachers do not facilitate practical work). Second, teachers commonly start a lesson by giving facts and then use an experiment to prove those facts (see for example class observations in Part I). Therefore, experiments do not test learners' constructs, but re-enforce textbook-held or the teachers' beliefs or facts. For example, an 'Experiment to prove Ohm's Law' is a common one, in science classrooms at Grade 12. Teachers also guide starting from simpler to processes that are more sophisticated, while rewarding learners' responses with marks and/or praise (Fosnot, 1996: 8; Winn, 1997; Conway, 1997: 1 – 2; Child, 1997: 10). Teachers are advised to avoid remarks that can dampen a learner's enthusiasm, such as 'failed' or corporal punishment. Science practical exercises comprise of worksheets by which learners adopt the ways of scientists,. (Again, let us note that the behaviourist association of stimuli and behaviour resembles the positivist principle of causality).

The changes in the curriculum during the early 1990s in SA re-interpreted behavioural objectives in terms of cognitive processes (both concrete and mental), and scientific 'specific outcomes', and learners' interests relating to intrinsic motivation were acknowledged. While this does not mean a direct shift to cognitivism, the science processes or science specific outcomes are also listed in Margenau (1974: 751) as cognitive processes. Therefore, the changes implied that learning science could be or was explained in terms of cognitive processes (Margenau, 1974: 751).

There are varieties of cognitive theories, but Bruner's and Piaget's theories seem to have gained currency in SA. For example, lower classes and lessons start with concrete or practical work before abstract concepts, which are introduced as learners grow up, in line with cognitive development. This has sometimes led to dealing with a single concept in different Grades – for example, Ohm's Law appears in Grade 8 and 12, which teachers complain about.

As stated earlier, the DoE recommends practical work, group discussions, and introducing a lesson by prior knowledge following Piaget (as cited in Driver *et al.*, 1994: 7, and Scott, Dysin & Gater, 1987: 7), who stated that schema are constructed upon physical and social interactions, and interpretations of those events, as well as prior knowledge. Furthermore, the DoE's demand for practical work is recognition of Adey's (1987: 17-19) advice for science skills such as control and exclusion of variables, frames of reference, and ratio and proportion.

Another important measure in the curriculum is the recommendation for project work that focuses on real problems facing learners in their home environments. For example, learners are lately involved in projects on saving energy and the environment. Such projects involve situated cognition because projects increase relevance of knowledge and of processes (Anderson, Reder, & Simon, 1996) if the focus is on problems in learners' environments. Furthermore, it is assumed that this situation-specific learning enhances self-concept and a feeling of control over one's own success (Weiner cited in Rieber, 1992: 99), and increases learners' abilities to apply knowledge in appropriate conditions, thus fostering invention and creativity.

Cognitive load is likely in practical and project experiences, but there are recommendations for reducing it. For example, the South African curriculum advises teachers to deal with a limited number of concepts or outcomes at a time.

Problems teachers face against implementing true cognitive classroom approaches include a shortage of tools, resources, and the limitations imposed by time and lengthy syllabi, which reduce the freedom of a learner to explore and acquire knowledge.

Constructivism in the learning of science

Many theorists acknowledge the importance of constructivism to science learning (E.g., Bodner, 1986; Campbell, 1998; Tsai, 2000;). Both cognitive and social constructivism are currently in SA desired classroom approaches in the new curriculum (C 2005), although this is not explicitly stated, and the two ECPs would be more beneficial if they could encourage constructivism in science classrooms.

As with cognitivism, constructivism ought to start when a learner is faced with problems s/he can solve using process-oriented hands-on activities and some objects (Kumar, 1994: 59; Dede as cited in Dede, Salzman, Loftin, & Ash, 1997; Scott, *et al.*, 1987: 7; Driver *et al.* ,1994; Campbell, 1998; Tamir, 1996: 109; Harlen, 1993: 28-36; 2000), in contexts to which the learnt knowledge is relevant (Duffy & Cunningham, 2001: 179). Learners sometimes get a chance to manipulate and transform objects (Driver *et al.,* 1994: 6; Mwamwenda, 1993: 71) during practical work in laboratories (Bodner, 1986) or through projects in their school and home environments (Hein & Lee, 2000: 1). Additional cognitive and constructivist measures include continuous assessment (Yore, 2001).

But in many South African classrooms, lessons are teacher-centred, and examinations driven, such that learners are rarely provoked to critically challenge their conceptual schemes as recommended in Geelan (2000: 4) and Yumuk (2002: 142).

Conclusion on learning strategies - No single theory explains all learning

The debate above on learning theories informs designers of instruction that all learning theories are essential in a programme – that is, learning theories are complementary, and each explains some aspects of learning. For example, Burton, Moore, & Magliaro (2001: 65) conclude that behaviourism can account for situated cognition and social constructivism. Ertmer & Newby (1993) recommend choosing a learning theory based on how much the learner knows. Thus, for example, a behavioural approach can be effective for knowing what; a cognitive strategy may be adequate for solving problems in unfamiliar situations if the learner knows the concepts involved and rules (knowing how); and constructivist strategies could be suitable for clarifying problems. Furthermore, Herron (as cited in Bodner, 1986: 873) argues that it is normal for people to start from the concrete to abstract whenever they encounter new experiences or knowledge.

Learning science, and important aspects of learning such as *motivation, lesson structure, sequence of concepts, and reinforcement* cited in Sprinthall & Sprinthall

(1990) indicate an apparent intercourse between cognitivism, and behaviourism. Jonassen *et al.* (2003: 2-9) capture this intercourse in providing definitions of learning that represent each learning theory, and in summarising the intercourse into five attributes of learning: active (manipulative and observant), constructive (articulative and reflective), intentional (reflective and regulatory), authentic (complex and contextual), and cooperative (collaborative and conversational) learning, which, they argue, computer technology can provide. Atwater (1996: 831) alludes to this intercourse in the statement that *no one epistemology can serve to explain what happens in science learning and teaching.*

ASSESSMENT

Introduction

Assessment appears in two ways in this study. In the first one, a CAA programme was evaluated for its diagnostic value. In the second case, assessment was one of the methods of gathering data for evaluating the two ECPs (Fraser, 1991: 2-3; Madaus & Kellaghan, 1992: 119-120).

The NRF (2004) believes that assessment is among the most important issues that require attention in schools, because it contributes to poor learning. Problems in assessment include the use of invalid tasks that are insensitive to the kind of learning or teaching applied in class, and of the nature of a subject.

The use of low cognition Multiple-Choice Questions (MCQs) is one possible source of problems. Objective-type testing has escalated (King & van den Berg, 1992: 23) especially in science (Madaus & Kellaghan, 1992: 127). For example, recall and simple calculations MCQs contribute approximately 30% to 40% of the total Matric science marks in the South African Matric examinations (my experience as examiner and marker). It is thus important to train teachers in setting higher order diagnostic MCQs, especially in light of the fact that most of the computer software available in SA use MCQs (The Department of Computer-Based Education, University of Cape Town, 2000: 2; Tamir, 1996: 96-98).

Assessment and evaluation

Assessment and evaluation as synonyms

Evaluation and assessment are sometimes used interchangeably even in professional literature that it is necessary to highlight some aspects of their relationship. Several sources show them as synonyms, for example: Eisner (1979); Lawton & Gordon (1996); The Maricopa Centre for Learning & Instruction (2000); and Harlen (2000). According to Percival & Ellington (1984: 100) as well as Madaus & Kellaghan (1992: 119), one reason for the synonymous use of assessment and evaluation is lack of agreed-upon meanings.

The DoE's definition is more comprehensive and relevant for this study, since the study was conducted in SA. The DoE defines assessment as the process of

> *identifying, gathering, and interpreting information about a learner's achievement, as measured against nationally agreed outcomes for a particular phase of learning. It involves four steps: generating and collecting evidence of achievement, evaluating this evidence against the outcomes, recording the findings of this evaluation, and using this information to assist the learner's development, ...* (DoE, 1998: 3).

The DoE's definition makes evaluation a part of assessment.

However, in this study, assessment and evaluation are considered different concepts as is the case in Fraser (1991), Percival & Ellington (1984: 100), Madaus & Kellaghan (1992: 120), Elliot (1991: 217), as well as Lloyd-Jones, *et al.* (1986: 1). Furthermore, I look at assessment and evaluation in the context of school curricula. Generally, assessment emphasises the gathering of data, while evaluation focuses on the interpretation or use of that data (Fraser, 1991: 2). I show the differences between assessment and evaluation in the following literature review starting with evaluation.

Evaluation

There are many purposes of evaluation listed, for example, by Weiss (1998: 5), but according to Harlen, (1980: 57), all of these purposes relate with the concern about pupil performance and new curriculum materials. Similarly, Lloyd-Jones *et al.* (1986: 1) articulate that evaluation broadly questions the worth of a course or activity or program in relation to the intended outcomes. Therefore, evaluation is important for making decisions on curricula issues, which include *context, general aims ..., and curriculum materials (*including *software)* (Madaus & Kellaghan, 1992: 128). While assessment can be used as a source of data for evaluation (Percival & Ellington, 1984: 118-119), it would be fair to evaluate the validity as well as the value of those assessment methods, techniques, and the data they produce. Madaus & Kellaghan (1992: 120), and Weiss (1998: 5) advise evaluation against an objective.

I review evaluation specifically in the context of ID later in this thesis.

Assessment

Traditional South African curricula took assessment to mean the measurement of the quantity of knowledge that a learner has. This is a definition of assessment that is popular in literature such as in Madaus & Kellaghan (1992: 120). The other definitions appear to go beyond knowledge. For example, assessment has recently included other performances (The National Center for Research on Evaluation, Standards, and Learner Testing, 1999). In case of science, these might include practical work, in which there are numerous processes and skills.

In this study, I use the DoE's (1998: 3) definition of assessment without making evaluation a subset of assessment. Thus, assessment is the process of identifying, gathering, and interpreting information about a learner's achievement, as measured against South African science outcomes for a particular phase of learning.

Assessment and testing

Until recently, many educators in South Africa equated assessment with testing. For example, continuous assessment implied continuous testing to some educators. It is C2005 that has highlighted the differences between assessment and testing. Literature too has shown that trend. For example, Hein & Lee (2000: 2) state that *assessment is a more modern and more inclusive term, than traditional 'tes*ting'. However, Fraser's (1991: 2) definition of a test, which is *a set of items structured in a specific way and applied with the intention to measure a given attribute among a selected sample of the population*, seems adequate for this research. Therefore, a test is one of the methods of assessment (Harlen, 1993: 158). A 'test' is traditionally part of teaching and learning (Madaus & Kellaghan, 1992: 126).

Assessment and learning

Few teachers in SA would clearly relate assessment to the learning strategy they apply. One of the reasons for this is that the DoE does not explicitly recommend a learning theory, and it does not identify a learning theory upon which it bases its assessment policy.

However, it seems difficult to arrive at possible solutions to some of the problems in science classrooms without an understanding of how learners are assessed, and therefore in what paradigm they are assessed. That is, that understanding requires the knowledge of how assessment relates to learning strategies. Indeed, it will be established that one of the problems in science classrooms might emanate from applying assessment forms that are discrepant with the learning strategies in classrooms.

This is because; assessment can be used to check on the learning process with the aim of identifying what requires improvement (The Maricopa Center for Learning & Instruction, 2000; Tamir, 1996: 94). Assessment supports learning (Gipps, 1996: 251) by providing evidence of what is taking place in class (Little & Wolf, 1996: ix), or should be an integral part of the curriculum (Madaus & Kellaghan, 1992: 126), and so should be informed and validated by the philosophies of learning adopted to the extent that the kind of learning intended, determines the form of assessment used (Gipps, 1996: 251-252; Ertmer & Newby, 1993). In fact, assessment appears to have undergone changes in concert with learning theories (Madaus & Kellaghan, 1992; Gipps, 1996), as can be established below.

Behavioural assessment assumes that learning comprises basic skills that are acquired through rehearsals of what was taught in class or in the textbook, and that it is observable and/or measurable (Birenbaum, 1996: 5; Cunningham, 1991: 13), since behaviourism takes learning to be a change in behaviour. Behavioural assessment relies on psychometric analyses (Birenbaum, 1996: 3), and deals with one attribute or concept per item, and an item's mark correlates positively with the total score, such that discrepant items (items that do not correlate highly with the total score) are removed (Gipps, 1996: 254). Thus, it comprises tasks whose answers are *correct or incorrect, right or wrong* (Scott *et al.*, 1987: 19). Such assessment assumes that there can be consensus on educational goals and objectives. Debate and essays would be marked against clearly set out criteria. So it comprises predominantly low cognitive tasks.

Data from behavioural assessment can easily be analysed and has an expected 'normal' distribution (Birenbaum, 1996: 5; Gipps, 1996: 252-253) (i.e., can be norm referenced). Similar marks and statistics mean the same thing, and so, behavioural results can be compared. This is possible because behavioural assessment checks whether a learner knows the content (Fuchs, 1995: 1-2; Gipps, 1996).

Therefore, it is reasonable to assume that behavioural assessment is objective (Hannafin, Kim, & Kim, 2004: 13); i.e., it is product-oriented and yields results that can be evaluated against pre-set curricular objectives, and targets (such as pass rates) agreed upon by the stakeholders. Such attributes seem to make behavioural assessment desirable in SA, where stakeholders are interested in immediate and reliable feedbacks for accountability.

However, accountability also requires that experts process assessment, instead of the teachers who actuate the instructional process (Birenbaum, 1996: 5; Madaus & Kellaghan, 1992: 125). This separation raises serious doubts regarding the validity of assessment, as it may not directly relate to activities in class. Hence, behavioural assessment is rule-bound, allocates time limits, and assesses the product, with no regard to the processes of learning (Birenbaum, 1996: 6). Furthermore, Fuchs (1995: 2) complains that *behavioural discrete tasks do not necessarily add up to important and applicable outcomes,* and Birenbaum (1996: 5) points out that *behavioural assessment can encourage teaching to the test or even teaching the test.* That is, the assessment system dictates a behavioural approach to instruction, which can limit the teacher's instructional options. This is the kind of assessment that Tamir (1996: 95) argues leads teachers to focus on completing the syllabus and, therefore to rote learning. Therefore, results from behavioural assessment might only be useful for diagnosing memory, but not misconceptions.

Cognitive assessment is in concert with behavioural assessment in dealing with small amounts of knowledge at a time, and on assessing simpler before more higher order and abstract concepts. Beyond this, cognitive assessment seeks understanding

(Madus & Kellaghan, 1992: 127). Assessing levels of understanding entails finding out 'how well' rather than 'how many', and:

- *Deals with the individual's achievement relative to himself rather than to others (i.e. criterion-referenced)*
- *Takes place in relatively uncontrolled, and unregulated conditions and so does not produce 'well-behaved' data*
- *Embodies a constructive outlook on assessment where the aim is to help rather than sentence the individual* (Wood as cited in Gipps, 1996: 255)

Similar advice appears in Jonassen *et al.* (2003: 228). Furthermore, cognitive assessment would consider the stages of intellectual development based upon Piaget's cognitive theory, such that tasks for younger learners ought to be easier and more concrete.

Mwamwenda's (1993) inferences of possible differences in intellectual (i.e., cognitive) development in different environments and cultures imply that cognitive assessment tasks might not be universally valid. This necessitates the consideration of Salviati's (as cited in Cunningham, 1991: 15) argument that, assessment should be embedded in situations that a learner has experienced and where it arises naturally. That is, assessment should be specific to a context and individual. Furthermore, Piaget's cognitivism might be problematic in that, at least for me, cognition might not be as linear as the whole experiences I have received in life, which include simple and complex phenomena simultaneously.

In response to the need for a more individualised assessment, constructivist assessment seeks a more detailed or comprehensive cognitive assessment. Constructivist assessment that is more subjective to the learner's cognitive processes (Hannafin *et al.,* 2004: 13), and environment and is integrated into a lesson such that it captures the interpretations and constructions a learner is making (Pachler & Byrom, 1999: 126; Natal College of Education, 1997: 109; Birenbaum, 1996: 6 – 7). Learners construct a response rather than simply choosing from pre-selected answers (Jonassen *et al.,* 2003: 228).

Constructive assessment is a basis for continuous and diagnostic assessment and suits practical work in science, for which Ryan & DeMark (2002: 67) advise constructed-response items. Constructive assessment seeks learners' constructs, especially in real situations (Cunningham, 1991: 16) or is about the learner's lived experiences. Additionally, constructivist assessment involves the learner in the process and consists of a variety of methods that make data more authentic. For example, one way by which a teacher can assess or check the problems or conceptual level of a learner is by use of think-aloud protocols or dialogue with the learner – i.e., asking the learner to say what one is doing and why one is doing it. The learner can then reflect upon what s/he has already done. Reflection is like articulation, except it is pointed backwards to past tasks (Bereiter & Scardamalia, 1989).

This makes constructivist assessment complex. Jonassen *et al.* (2003: 229) advise that such complex assessment requires a rubric, which they define as *a code, or a set of codes, designed to govern action.* They go on to clarify that *rubrics have taken the form of a scale or sets of scales* used to assess complex performance, and that learners could participate in developing a rubric. Jonassen *et al.* (2003: 230-232) identify characteristics of an effective rubric as including: important elements; well-defined elements; distinct, comprehensive, and descriptive ratings that cover the range of expected performances; provides rich information about multiple aspects of the performance.

However, this kind of constructivist assessment might look informal, which Hickey, Kindfield, Horwitz, & Christie (2003: 529) argue might de-motivate learners and be less effective.

Diagnostic assessment

Why diagnostic assessment?

Teaching would be difficult without mechanisms of checking the understanding or problems learners have. Hence, diagnostic assessment happens frequently in class, but is not formally applied in science tests in SA.

Several theorists (Lawton & Gordon, 1996: 88; Bright, 1987: 71-83; Fuchs, 1995: 1; National Council of Teachers of Mathematics, 1995: 3, 5; Hein & Lee, 2000: 3; Taiwo, 1995: 3; Linn, 2002: 40; Gipps, 1996; Harlen, 2000; Pollitt, 1990: 879; Black, 1986: 13-16; Fraser, 1991: 5-8; Sanders & Mokuku, 1994; DoE, 1996, 2000: 34; Little & Wolf, 1996: xi) give numerous reasons for diagnostic assessment. Among the reasons for diagnostic assessment are: identifying and analysing specific abilities, difficulties, and incorrect conceptions. That is, diagnostic assessment is useful for scrutinising learning difficulties, so that appropriate remedial help and guidance can be provided. Therefore, diagnosis is more effective at an individual learner level and when focussed upon a specific area of knowledge or skill, but is difficult to apply in classrooms with large numbers of learners and lengthy syllabi.

Diagnostic assessment in class

The DoE (1998) explains that diagnostic assessment takes on a truly supportive and formative role by guiding the learner and by helping the teacher to plan appropriate activities to meet the learner's needs. This is possible when diagnosis is part of the *teaching sequences*, which accommodate the *re-teaching act* (Bright, 1987: 81). This is easier with continuous and constructivist assessment because the constructs, processes of science or outcomes learners achieve are continuously diagnosed, improved upon and remedied continuously.

Errors and mistakes

Answers can contain errors or mistakes, and the difference between the two is important in diagnosis. Bright (1987: 72) explains that a mistake is an incorrect answer while an error is a consistent pattern of mistakes made in response to a series of similar exercises or questions. Bright advises that diagnosis should determine and correct the errors, not mistakes, and explains that the difficulty in remediation of errors is that some errors can lead to correct answers. Therefore, repeated exercises are necessary to reveal errors in understanding. Furthermore, the teacher should check the ability of learners who give the same wrong answers; especially to check whether weak learners get difficult items right, or able learners get ordinary questions wrong (Pollitt, 1990: 879). A repeated error might indicate *inter alia* a teaching fault rather than a learning fault (Pollitt, 1990: 885). Hence, Pollitt (1990: 877) recommends that item banks comprising different levels of difficulty on each topic should be set so that diagnostic data can relate ability to kind of task.

Validity of test items

General validity

Test items have to be valid in order to assess accurately. Test item *validation* is the process of evaluating the degree to which theory and evidence support a specific test-score interpretation or use (Haladyna, 2002: 94). Therefore, validity is a central and an important consideration in evaluating the appropriateness of all forms of assessment, and must guide the development, application, and the use of results of assessment (Ryan & DeMark, 2002: 67; Linn, 2002: 27). Therefore, I had to validate test items to avoid role conflicts (Little & Wolf, 1996: xiii; Linn, 2002): in this case, the role was diagnosing learners' problems in understanding science. However, there is disagreement over the chronology and kinds of validities, to the extent that some definitions can cover multiple validities (E.g., Salvia & Ysseldyke, 1988: 132; Linn, 2002: 28-33). Hence, Linn (2002: 46) advises for prioritising and addressing the most critical validity questions. I thought that the following validities were most important in this research.

Construct validity

The term 'construct' is a non-observable parameter, which accounts for regularities or relationships between traits (Taiwo, 1995: 8; Gay & Airasian, 2000: 167-168; Hargis, 1995: 154). According to Hargis (1995: 154), *constructs are difficult to define precisely* because they are often *theoretical or hypothetical*, but include motivation, sociability, intelligence, and interest in something. An acceptable definition of construct validity is the extent to which a test measures an intended characteristic or construct (Salvia & Ysseldyke, 1988: 139; Taiwo, 1995: 8; Gay & Airasian, 2000: 169). Construct validity ensures that the test measures what it is set out to test *and not something else* (Lloyd-Jones *et al.*, 1986: 36). Therefore, *a clear understanding of what is to be measured* must be established before setting the test (Salvia & Ysseldyke, 1988: 132).

Construct validity is determined by correlating measures of observable criteria that the construct is highly related to (Taiwo, 1995: 8). That is, construct validity is determined on *indirect evidence and inference* from conducting experiments to demonstrate that the *test is not a valid measure of the trait or construct* (Salvia & Ysseldyke, 1988: 139). For example, although intelligence tests could predict achievement, the correlation between a test score on intelligence and a test score on achievement does not mean that the test on intelligence measures achievement, because there could be other factors contributing to achievement. One has to investigate other factors that could affect achievement; when these other factors do not correlate with achievement, it can be deduced that intelligence tests measure achievement.

Participants were asked to state whether they thought that the test items were in fact about science and not something else.

Content validity

Content validity is deemed the most important among validities (Hargis, 1995: 150, 156). Content validity is the degree to which a test measures intended content or outcomes, and is the means of checking whether the content was taught, belongs to the syllabus, and is of the expected level for that Grade (Gay & Airasian, 2000: 163; Fraser, 1991: 19; Taiwo, 1995: 7; Salvia & Ysseldyke, 1988: 133-134; Hargis, 1995: 150). Content validity is alternatively known as 'curricular validity' (Taiwo, 1995: 6). Lloyd-Jones *et al.* (1986: 36) refer to curricular validity as the justification of the objectives of the course or lesson unit that is being tested. Another aspect of content validity is whether the test is in concord with teaching methods (Lloyd-Jones *et al.*, 1986: 38).

Besides the cognitive domain, content might include attitudes, which, according to Gagné (1985: 219-242) and Harlen (1993: 37-44), is the tendency of somebody doing something. White & Tisher (1986: 892) state that Likert tests are commonly used to measure attitudes towards science. However, Harlen (1993: 189) reports that attitudes towards science can directly be assessed, giving an example of assessing *the willingness to change ideas* given evidence which contradicts the usual knowledge, or to challenge a conclusion given *insufficient evidence*, or assessing *respect for evidence*. A question to check on content during the preliminary survey was: How accurately does this test represent the content of the curriculum?

Face validity (Fairness)

Sanders & Mokuku (1994: 482), Fraser (1991: 20), and Gay & Airasian (2000: 164) show that the meaning attached to face validity is confusing. The confusion is apparent in the following wide range of explanations about face validity. Face validity is related to content validity in that it is *the extent to which a test appears to measure what it claims to measure* such that face validity might be an initial stage in establishing content validity (Gay & Airasian, 2000: 164). Taiwo (1995: 7) clarifies that face validity *is the reasonableness* concerning *the background of the testees,* and concerns *the relevance, adequacy, and coverage of items.* Taiwo states that face validity is established by questions like, *how do the test items look like in the light of the objectives of the test?*

To the DoE (1998), fairness means that an assessment is not biased, and *offers an equal opportunity to learners of both genders and all background.* Thus, content, context, and performance expectations of a task should reflect knowledge, values, and experiences that are equally familiar and appropriate to all learners; tap knowledge and skills that all learners have had adequate time to acquire; be as free as possible of cultural, ethnic, and gender stereotypes or attitudes, beliefs, or values. For example, the context including football becomes a biasing factor if particular groups of learners know less about football than other groups of learners. Language is another biasing factor. Markers should also be free of preconceptions about the abilities of different learners. A question to test fairness and face validity can be: Is the test fair?

Factors that affect validity

There are numerous factors, which have to be considered because they affect validity. Among those in Salvia & Ysseldyke's (1988: 109, 140-142) list, is reliability, because it is one of the advantages of using CAA. Reliability means *dependability, or trustworthiness* (Gay & Airasian, 2000: 169), *confidence* (Hargis, 1995: 143), or *consistency* (Dietel *et al.*, 1991: 2; Fraser, 1991: 35; Hargis, 1995: 143; Lloyd-Jones *et al.*, 1986: 39) of an assessment. A reliable assessment gives the same results if it is done again or if re-marked (Northwest Regional Educational Laboratory, 2000). CAA eliminates errors in processing results and so increases reliability. Unfortunately, stringent adherence to the requirements for reliability can affect validity negatively. Taiwo (1995: 9) warns that *an instrument could be reliable to the extent that it gives the same measure, and yet not valid in that it could consistently give a wrong measure.*

Multiple Choice Questions (MCQs)

MCQs

As was stated in the introduction, tests or examinations comprise significant proportions of MCQs, to the extent that the marks a learner obtains from MCQ might decide whether s/he passes or fails. Therefore, it is important that teachers learn how to set valid MCQ, and learners practice these. The Department of Computer-Based Education, University of Cape Town (2000, section 2.1) defines a Multiple Choice Question (MCQ) as a question in which [learners] are asked to select one alternative from a given list of alternatives in response to a 'question stem'.

There are a wide variety of classifying questions (E.g., King & van den Berg, 1992: 22-24; Twomey & Miller, 1996: 5). Generally, subjective items are those in which answers and marking are subject to the individual examinee and examiner, while objective tasks are those in which answers are definite and therefore can be marked with a high degree of reliability. MCQs are commonly objective and rarely subjective, since choices cannot exhaust possible subjective thoughts, and are therefore likely to be behaviourist. Thus, it is challenging for a teacher to set diagnostic higher order or constructivist MCQs.

Advantages of MCQs

The popularity of MCQs arises from their advantages. King & van den Berg (1992: 22-24) list numerous advantages of MCQs, which make them attractive for use. These include the following:

a. Assessing very large numbers of candidates;
b. Reducing problems due to language, since answers are normally either provided or short;
c. MCQs are easier to incorporate into CAA, and to analyse statistically
d. Offering the possibility of dealing with a wider range of topics and cognitive levels in a short time; and
e. Easier and accurate marking, as well as administration (E.g., University of Cape Town, 2000: section 2.2; Tamir, 1996: 96).

General principles of setting MCQs

Teacher's courses include skills of setting MCQs. Bright (1987), Croft *et al.* (2001: 58), Dreckmeyr (1991: 50-76), as well as Twomey & Miller (1996: 6) outline some of the general principles, which we applied in setting tests in this investigation. For example, tests comprised the subject matter that had been taught, used language of the appropriate reading level, avoided tricky questions, and were deemed to be fair to learners. Other important principles that we followed to design MCQs included:

- Each item dealt with one clearly stated problem;

- Clear, correct, and simple language was used;
- Short stems were preferred;
- Negative stems were avoided, and highlighted if used;
- The correct alternatives were as much as possible made to be clearly correct or clearly the best; but distracters were attractive
- Verbal clues; for example by different lengths of alternatives, were avoided
- A uniform format was used

MCQs have traditionally been objective or behavioural 'correct – incorrect' or lower cognitive levels such as recall items. Such items do not assess understanding. MCQs can test understanding if, according to Tamir (1996: 96-97), the design shifts to *'correct – best answer'* formats, which encourage thinking and *a wider range of cognitive abilities*. Tamir explains that distracters in a *'correct – best answer'* contain some *factually correct information*, forcing the learner to *analyse the various options*, and can function like *a Piagetian classical interview in which the interviewer is not fully satisfied even with the correct answer* given. Such MCQ items could have a high diagnostic potential.

Construction of diagnostic multiple-choice items

Setting diagnostic tasks is commonly easy in open-ended essay type of questions. It is not easy to set diagnostic MCQs. Nevertheless, to achieve the 'correct – best answer distracters, a number of questions should be set on one aspect or concept at a time. Bright (1987: 71-83), Tamir (1996: 97, 107), Maloney (1987: 510-513), Amir & Tamir (1994: 94-95), and Fraser (1991: 5-8) give guidelines on setting diagnostic items, some of which the teachers and I used as follows:

- Using known misconceptions as distracters
- Using learners' answers to open-ended questions and processes a learner might use for constructing distracters
- Ranking items and giving reasons for the ranking - candidates to defend their reasoning and understanding
- Use of paired-problem-solving activities, in which for example, a concept is needed before calculations.
- Penalising for ludicrous choices (choices that show complete misunderstanding) by awarding a negative mark
- In 'confidence in chosen response', learners are asked to choose the best answer and to indicate if they are sure or not sure of their choice, and then

to choose the second best answer and indicate again whether they are sure or not sure. The following marking procedure is an example of this method: Correct sure = 2 points; Correct not sure = 1 point; Incorrect not sure = 0.5 point; Incorrect sure = 0 point

- Ask learners to give justification for their choice
- Provide data and ask learners to describe (analyse, synthesise, and evaluate) the situation that data represents.

It appears from the above that diagnostic MCQs might come close to rubrics. Fraser (1991: 5-8) points out that different subjects, and possibly curricula, might require different strategies.

PLAY

Playing and learning strategies

Playing games is one activity during which a mixture of learning theories (in various proportions) can be applied. Furthermore, Sugar & Sugar (2002: 4-8) show how playing games supports multiple intelligences as well as experiential learning. Play is a mixture of learning strategies (Table 9). I explain these strategies in more detail in the next section.

Table 9: Possible applications of learning theories in playing games

Desired attribute	Learning Theory	Application
Motivation: Malone & Lepper (1987), Ayayee & Sanders (1998: 53, 56), Draper (2000), etc.	Behaviourism = extrinsic: E.g., winning and scoring. Cognitivism = intrinsic: E.g., exploration, control, fantasy, and imagination. Constructivism = intrinsic: E.g., manipulation and constructing models, and creativity.	Playing challenging games, with music, scoring, and open micro worlds
Scaffolding and helping disequilibria and transformation: Rieber (1996a); Duffy & Cunningham (2001: 183).	Cognitive apprenticeship Social constructivism	Provide effective and immediate feedback in a game.
Experiential learning: Kraft & Sakofs (1988) and Adey (1987) Learning by doing cheaply (without contravening ethical rules)	Cognitivism = schema reorganisation Constructivism = Schema construction	Simulate real life situations and activities in games.

Play can cover the whole classroom space (Table 10). For example, games can be Socratic or constructivist (Laurillard, 2000), and playing games seems to achieve different kinds of motivations explained by different learning theories (Bindra, 1969).

Playing games

Play is not easy to define because it is subject to individuals, time, context, and culture. Play is human nature, but is surprisingly not used in learning science, or is used only in lower classes. There are attempts in SA now to increase the use of play in higher classes to learn science.

The times when a person attaches emotion to a task can be described as play (Rieber, 1996a; Draper, 2000). The process and emotions involved in play is flow. According to Rieber (1996a), Quinn (1997), and Draper (2000), the term "flow" or "autotelic experience" originates from Csikszentmihalyi & Csikszentmihalyi (1988) who explain that flow is the experience of extreme happiness, enjoyment, and satisfaction to the extent that a person flows along spontaneously with the activity. Flow involves some level of active, often physical, engagement, and processes. There are two kinds of flow:

> *"U-flow" is a unconsciously managed flow of actions without being able to remember anything that happened along the way.*
>
> *"C-flow" is a flow of actions that is managed by and fills the consciousness of the actor. C-flow may not involve physical actions, but always involves complete mental attention. C-flow is a balance between boredom (nothing seems important) and anxiety (when too many goals and actions seem important, urgent, and uncertain to be satisfied).*

Rieber (1996a) and Draper (2000) note levels of flow (E.g., participation for fun, problem solving [development of physical and mental perceiving tools], and catalytic action (intuitive, spontaneous, and creative action). The first part is a sort of bait that lures a learner into the programme. The later levels lead to learning (Rieber, Smith, & Noah, 1998).

One can play a variety of things, but Instructional Design (ID) has focussed on using playing games to enhance learning. Similar to play, the concept 'game' has long been indefinable (Quinn, 1997). In some instances 'play' seems synonymous with 'game'. However, the differentiation between play and game could be made on the basis of whether there is a play that is not a game. One can play a tennis or soccer game. One can also play drama, which is in some cases a simulation of real life situations.

So a play is not necessarily a game, and while a game is real, a play can be a simulation. In that case all games are played, but not all playing involves games. Some games are serious, whilst others are for fun. On the other hand, play is apparently always for fun.

Therefore, Horn & Cleaves's (1980) notion of a game is meaningful and relates with a definition of a game that is adopted in this thesis: A game is play constrained by a set of explicit rules particular to that game and by a pre-determined end-point.

There are cooperative and non-cooperative games, according to Levine (2001), but I think players can choose a game to fall in either of the two categories. For example, players can decide to cooperate in a game of cards or to compete (be non cooperative). Furthermore, since games appear to be subsets of play, characteristics of play, such as the subjectivity, apply to games.

Games offer a practical means of meeting the microworld assumption of self-regulation, and offer many intriguing psychological and social insights to microworld design (Rieber, 1996a: 49-50). Rieber sees the use of games as that of attracting people to knowledge, through fantasy, challenge, and curiosity. The British Educational Communication and Technology Agency (BECTA) (2001: 1) explain some of the reasons for these assumptions. BECTA argue that *games use technology to represent reality or to embody fantasy,* and *provide an environment in which action can be practised or rehearsed with, ultimately, little consequence. Furthermore, some games can be cooperative and non-cooperative* (Levine, 2001, BECTA, 2003).

A designer has to plan a game in such a way that a player achieves flow. Rieber (1996a) and Draper (2000) give factors that a designer has to consider. For example, the game must have clear goals and the player has to see that s/he is in control to the extent that those goals are achievable. Another important element is that a game provides immediate, clear and consistent feedback as to whether one is reaching the goals. The feedback as well a well designed challenge motivate a player to concentrate effortlessly and so to be absorbed so that time passes without notice.

Play and learning

Problem solving and catalytic action happen as a player discovers the outcome of a process (the consequences of some rules), or "learning by exploration". (E.g. will I win? Can I build this chair? And if so, how?) (Draper, 2000). As such, play is uncertain, and reaching the destiny involves the player making decisions in order to solve puzzles and problems along the way, which call upon cognitive and intuitive speculation. Thus, effective games are often good in the sophistication of the user interface and/or content (BECTA, 2001: 3).

Sophisticated games require players to use logic, memory, problem solving and critical thinking skills, visualisation and discovery (BECTA, 2001: 3). In this way, play leads to the development of mental and intellectual abilities such as 'organizational strategies' and interactions that lead to focused learning (paying attention), problem-solving strategies, practical reasoning skills, retention and memory strategies (grouping, imagery, and structured review), motivation, social development, and compensatory strategies (guessing meaning intelligently) (Birenbaum, 1982: 4; Rieber, 1996a; Rieber, *et al.*, 1998; Kirby, as cited in Mosimege, 1997: 530; Hogle, 1996: 11). Games also improve affective strategies (anxiety reduction and self-encouragement) (Hogle, 1996:11).

The benefits accrued from learning associated with playing games based on constructivist-teaching environments have been articulated widely (E.g., Hogle, 1996; Rieber, 1996a: 46; Amory, 1997; Turoff, 1995; BECTA, 2001). Exploration thorough games in virtual environments (VEs) provides constructivist opportunities for building, and for changing concepts (Rieber, 1996a: 45), because it calls for a high degree of metacognitive activity and self-initiative to master unstructured situations, rules (or generalisations), and discoveries, and to exercise the relationships between these and their consequences (Leutner, 1993: 114; Winn, 1997). Hence, exploration of, and interaction with, a game during play in constructivist virtual environments can be similar to a scientific investigation. *Zeltzer uses "interaction" to mean the extent to which the participant logically follows the laws that govern the environment* (Winn, 1997). An activity such as play that enables interaction, intrinsically "engages", and leads the learner through problem solving experiences (Quinn, 1997).

Self-regulated learning

Some of the compelling reasons for recommending computer games in learning are motivation and self-regulated learning within a constructivist framework (Rieber *et al.*, 1998). Rieber explains self-regulation as a process through which one resolves issues and obtains solutions. Zimmerman (as cited in Rieber, 1996a: 47) explains that self-regulated learning includes metacognitive activity (E.g., planning, goal setting, monitoring, and self-evaluation), and behavioural activity (E.g., selecting and structuring the environment for one's learning style). These processes fit Laurillard's (2000) explanation of a learner's active narrative construction. Play provides (narrative) frameworks and goals that require metacognition, and increases self-esteem when one scores or wins a game (Rieber, 1996a; Willis, 2000: 7). Self-regulated play (intrinsically motivated) can attract learners to science, and can inculcate responsibility for their learning and for outcomes.

Designing instruction for self-directed learning blends motivation with the learning process, and includes goals that are interesting for their own sake (Karaliotus, 1999). Such ID provides opportunities to learners to monitor their own learning, and provides to learners opportunities to alter their learning environment. Learners have the authority to learn what they value, to set goals, and to use learning methods they prefer (Karaliotus, 1999). But some support might be necessary, for example, according to Laurillard (2000), by providing for them conversational frameworks and defined task goals in narrative multimedia, which involve *inter alia* media controls. The Curriculum Initiatives Branch (CIB, 2002) recommends the inclusion of graphic organisers (mind maps), flow diagrams (sequence of ideas, procedures or events), sequence of illustrations (E.g., pictorial sequence of ideas, procedures or events), etc. These can be represented in three-dimensional qualitative frames. Exploration through these tools, along with interesting questions that elicit the learners' existing conceptual frameworks, and beliefs, create interest and stimulate curiosity in contexts that learners can relate to.

Motivation

Learners identify motivation as a major factor in learning (Ayayee & Sanders, 1998: 53, 56; Rieber, 1996a; preliminary survey). Play supports the two main forms of motivation. Draper (2000) asserts that extrinsic motivation refers to external reasons for action (E.g. working for pay). This explanation fits Hannafin & Rieber's (1989: 93) belief that extrinsic motivation is behavioural, since it is based on the nature of reinforcing stimuli. Draper (2000) states that intrinsic motivation refers to a person's inherent enjoyment in the activity for its own sake (E.g. eating, going to a movie), and explains that intrinsic motivation does not depend upon reward that lies outside the activity, but reward is in the successful termination of the activity or even in the activity itself", especially when the user possesses the power to control destiny. This is in concord with Hannafin & Rieber's (1989: 96-97) description of cognitive motivation – a motivation that depends upon an individual's felt need to engage into some activity for its value.

Bindra (1969: 11-12) explains and links extrinsic and intrinsic motivation: First, activities of the nervous system create desire. Second, a stimulus in the environment, such as food stimulates action. This stimulus might have affective properties, such as 'emotion'. Motivation is then a function of neural change, and its interaction with an external object. That is, motivational actions are naturally instigated internally, but are observed as a person acts to satisfy internal body needs. In other words, motivation is initially driven by the desire to participate in a task and is subsequently sustained by choosing to persist in the task (Karaliotus, 1999). Playing a game would follow a similar sequence: natural desire for enjoyment, and then seeking objects to play with. There is satisfaction through rewards and opportunities for further exploration or the player tries again or looks for another game. Thus, although self-regulated learning requires intrinsic motivation (Malone & Lepper, 1987), it survives on extrinsic rewards. Rewards can be highlighted through self-driven activities such as evaluation and monitoring, which (Hogle, 1996: 11) believes are characteristic of playing games.

Bindra (1969: 13) states that the generation and persistence of motivation is necessary for goal-directed actions such as exploration (which we have seen in the above includes cognition). Games are fun, enjoyable and motivating when a player overcomes challenges and solves puzzles (Draper, 2000). Malone & Lepper (1987) explain that challenge refers to the level of difficulty and to performance feedback for the player, and includes goals, predictability of outcome, and self-esteem. Malone also advises that games in which curiosity engages deeper cognitive processes are intrinsically motivating. Hence, learners reported more interest in games than in conventional lessons (Randel *et al.*, 1992: 268).

Application of play in a classroom

As alluded to above, playing games is traditionally an 'extra' curricula activity – that is, schools play after 'serious' study, often in the afternoon. Classes at lower levels, such as in Grade 1 and 2, involve some fun as a means of motivating learners, but play is not so prominent among adults where 'serious' study is desired. Introducing play into adult classrooms requires convincing teachers about the value of play, and then revising curricula. It might be better to introduce playing games in a lesson after learners have been exposed to the concepts in the game because pre-course knowledge was found to improve the utility of games especially for weaker learners for whom games would be valuable supplements to lectures (Randel *et al.*, 1992: 264).

Problems with play

Reports of the effectiveness of educational games, measured against learning have been inconsistent in different games and subjects. (Randel, *et al.*, 1992; Ivala, 1998). In the first place, some designers such as Rieber (personal interview, 2004), do not believe that games should be designed for a result.

First, this is because desiring a result and learning a process are not always compatible. Science is a process-oriented subject, while playing is a result-oriented activity. Hence, Rieber (1996a) argues that doing science is not playing because science is not necessarily done for a known result. Thus, the difficulty is to design games where each step of playing counts towards the final result, but at the same time the process contributes to understanding.

Following from the above, is to see how to use play without imposing it so 'seriously' upon learners (because imposing a game removes the fun of playing it). Quinn (1997) points out that a game is fun as perceived by the player – so it cannot be imposed. Draper (2000) explains that *not all computer game(s) give enjoyment (i.e. satisfy various kinds of intrinsic motivation)* because *motivation and so fun is not a property of an activity, but a relationship between that activity and the individual's goals at that moment.* This might be due to the subjective (individual and cultural) nature of enjoyment. For example, adding colour and music might not automatically add value to enjoyment. In similar light, *it is difficult to pitch games at the right level of interest and challenge for the user*, to the effect that *games may be too easy or too difficult to play, with a decrease in motivation in either case* (BECTA, 2001: 3).

According to Draper, what matters is *the demand level of a game – if it is to be fun, a game must be matched to the player's arousal level, which in part varies independently of the game, for instance with the time of day.* Designing and using games for education are complicated by the observation that teachers and designers of instruction consider motivation in terms of what they can do to get learners to study, and so motivation is often an "add-on" feature (Karaliotus, 1999).

The third problem is the observation that games are often gender specific (BECTA, 2001), with females taking on leisurely games. The problem of leisure extends to people's attitude towards play. Indeed, few 'serious' teachers would want a playful class. As such, lessons arising from play are not often desirable in the traditional educational curricula (Rieber, 1996a). Computer games present problems, which include inadequate directions, lack of options to bypass certain stages, and inadequate stimulus control (Gredler, 2001:531).

Finally, playing games might lead to addiction, and oversimplifying reality. Therefore, the implementation of games should be done with care and with a specific purpose in mind (Mosimege, 1997: 534). Games might also produce learners who have always got to be enticed to study. Mosimege recommends going beyond enjoyment and giving learners a thorough understanding of a game in class.

Conclusion on play

Although BECTA (2001) suggests ways of evaluating games, Quinn's (1997) observation that there are no systematic evaluation procedures should be kept in mind. It is apparent that there are pro and cons regarding the use of play in education, and the decision as to whether a game is beneficial or not is subject to the player and to a particular game. That is evaluating a game is likely to be idiosyncratic interpretative affair. Idiosyncrasies of constructions (for example due to differences between cultures) lead to an inability to communicate because there is no shared meaning (Duffy & Cunningham, 2001: 171). It is probably because of the subjectivity of play that it is difficult to obtain conclusive evidence about the use of games in classrooms. Research is needed to establish how to obtain the values teachers and learners attach to games in education, to know what kind of values teachers and learners attach to games, to see how games contribute towards solving problems in science classrooms, and to see how such games can be included in a curriculum.

My untested suggestions of a possible overlap between science and learning theories appear in Table 10. 'Class space' is representative of activities in a typical South African classroom (Table 10). There are some overlaps that appear, which pose challenges to the way we demarcate subjects or learning theories – for example, the overlap between Piagetian cognitive skills and science processes is a clear challenge for educators and instructional designers to decide whether their programmes are specifically for science or generally serve, in a cognitive way, all disciplines. A further complication is that some of these cognitive skills are also resident in behaviourism and constructivism, while science processes can be claimed by other disciplines. The questions are, should we then make learning subject-centred or integrate all subjects under learning strategies that we find acceptable? And what of assessment – is it fair to say a task is cognitive or constructivist or rather to claim that it is just a science question? These are questions beyond this research, which I think, require serious consideration.

Teachers in South African science classrooms do not seem to adhere to a single science philosophy or learning theory. Classroom practices (teaching and assessment) are mixtures of strategies, and can be found anywhere in the 'classroom' space (Table 10). Curricula have only tried to emphasise one or the other without completely implementing or eradicating any.
However, it appears that C2005 has moved science learning from the left upper area, diagonally towards the right-hand bottom of the table – i.e., towards constructivist and post positivist practices.

This reality informs instructional designers that a teacher's practice is subjective, and is difficult to locate. In other words each teacher might value an Educational Computer Programme (ECP) in relation to his/her way of teaching, and there are so many ways of teaching, probably as many as are teachers. One way to design an ECP is to get teachers involved in identifying their classroom problems as well as in evaluating ECPs against those problems and the recommendations in curricula. The other is to be guided by the curriculum; for example, C2005 recommends diagnostic continuous assessment and constructivism.

Table 10: Possibilities of overlaps between learning theories and science philosophy in South African science classrooms

		Learning theories		
		Behaviourism (Fundamental pedagogics) - teacher or instructor designs a learning process Learning = behavioural change. Desired behaviour re-enforced by extrinsic motivation. Combine a sequence of stimulus and response Causality	**Cognitivism** Learning = change in schema structure Concrete to abstract Age matters Experiential learning Multiple intelligence Situated cognition Cognitive load Cognitive conflict = conceptual change Spatial skills	**Constructivism** Learners are not blank and use prior knowledge and experiences to construct new meaning. They also help each other to agree on meaning
Science philosophy	**Rationalism** Knowledge deduced from reason, abstract innate ideas or prior knowledge (*a priori*), independent of sensory experience - reason and logic	Learners are given rules for solving contrived problems	Knowledge must be suited to age: E.g., toddlers are not expected to understand Ohm's law Schema constructed from logical deductions Science processes - mental	Learners find their own way of solving problems in their lives Debate (Philosophically – a way of constructing reality)
	Empiricism *Tabula rasa;* all knowledge is achieved a *posteriori* through our senses or observation	Learners follow worksheets No ideas from learners – learners fed with facts	Psychomotor skills are age-bound and experience matters. Apply relationships to solve current problems Science processes - concrete	E.g., Open environments that allow learners to test their constructs (Philosophically – a way of constructing reality)
	Positivism Precise, certain and objective measurements, and rejects subjectivity and human ideology, history and intervention or intuition Experiences must be validated	Rules to be followed are well laid out Data must be precise – no human error Must provide acceptable evidence	Science processes – mental and concrete	Triangulation = social constructs? (Philosophically – a way of constructing reality)
	Post positivism Individual experiences are valid realities Hypothesise-deduce Critical realism: Triangulation to reduce human error	**?**	**?**	*Interpretative(idiosyncratic)* *Negotiated meaning =socially constructed meaning* *Negotiated assessment*

Classroom space – play and assessment can belong anywhere in this space

PART III

POSSIBILITIES OF USING COMPUTER TECHNOLOGY IN EDUCATION

The basic relevance of technology, is in using it to solve problems, such as those discussed in the previous section. Evaluating computer programmes in education requires an understanding of how or why computer came to be used in education, and how it is being used.

Solving problems in science education using computer technology

Previous chapters have deliberated upon problems in science classrooms. Not withstanding the expenses of equipping schools with computers, software, educational programmes, and training necessary in schools, computer technology can offer some solutions to problems in science classrooms. Despite lack of evidence on the superiority of computers to other instructional media (Hannafin & Rieber, 1989: 91), I suggest possible solutions (Table 11).

Table 11: Possible ways by which computer-based instruction (CBI) could provide solutions to problems in education

Problem	Need	Possible contributions to solutions from CBI
Assessment, diagnosing, and remediation is difficult for large numbers of learners	Learners should assess themselves as frequently possible, while their work is marked and analysed immediately.	Computer-Assisted Assessment allows learners to work through tasks as many times as they wish, provides data on their performance immediately, which can be used for diagnosis and remediation.
Problems with the NOS	Application of selected philosophies	Computer-enhanced research
Learners do not have enough practical experiences and environments to construct knowledge	Provide open environments in which learners are not restricted to test their ideas	Constructivist virtual environments in which learners can try out their ideas freely
Science is a boring subject and the number of learners enrolling for it is dropping.	Find ways of making science interesting to learners	Make studying science interesting and enjoyable by use of playing scientific games
Drop in number of science education students	Make science education more interesting and accessible to	On-line interactive lessons
Restricted learning time because teachers become tired, and the timetable is crowded.	Provide alternative sources of knowledge besides the teacher and school	Educational software that is accessible anytime: E.g., on the school LAN and on the INTERNET
Science is a foreign culture	Get stakeholders involved in designing curricula	Design instruction that is context-sensitive, and involve stakeholders in its evaluation. Provide interactions that can modify the programme to suit the context and culture.
Some science processes are too long or very dangerous to learn practically in a school classroom or laboratory.	Find ways that can be used to simulate such processes	Computer-simulated processes

The NRF Focus Area Programme for 2005 identifies research in the use of computers in education as one of the focus areas of concern for SA. The NRF (2004) also notes the urgent need to train teachers, especially those who are denied the opportunity to use information technology.

In this section, I outline the popular models in Instructional Design and how these relate to learning theories. The models give an indication of the epistemological positions of designers, especially for ECPs that are placed in models. I believe that the evaluation of an ECP is more accurate and fair to the designer and potential end-users if the evaluator places it in a particular model.

Instructional design (ID)

I consider every action taken to design a teaching exercise as an instructional design. However, those who design learning using technology, especially in response to problems claim to own ID. On addition, ID is variously defined possibly because it requires different professionals, involves many stages and theories, and depends on how one views 'instruction'. It is difficult to pick out one definition, among those given for example by the Applied Research Laboratory (1996), and by Dick & Carey (1990). For the purposes of this research, I have composed one from the Applied Research Laboratory (1996), Reeves & Hedberg (2003: 119), and Shakeshaft (1999: 3), which I think focuses on the activities of this study. Instructional Design is a process involving the analysis of learning needs and goals and the development of an instructional programme to meet those needs, and the evaluation of the programme and process. This study is concerned with researching the evaluation and use of two ECPs in South African schools in light of learning strategies used in those schools.

ID offers alternative ways to solve problems (Berger, 1988) in education. Greening (1998) argues that technology in learning is likely to reduce crises in education, and to encourage constructivism.

Learning theories and Instructional Design (ID)

Introduction

It is important to understand the relationship between learning theories and ID because this relationship influences the strategies of evaluating ECPs, and provides light upon how an ECP could be integrated into a curriculum. It is thus expected that one aspect of ID that appears in its definitions in the above is its relationship with learning theories.

Wilson & Cole (1991) as well as Wilson (1995b) report opposing views and lack of clarity regarding the ties between learning theories and designing technologically enhanced instruction. Of course, there is simply too much on this topic to do justice in this research to the intricacies involved. This a simplified outline, touching on relevant aspects only.

Views on the relationship between learning theories and ID

The first view is that ID is independent of learning theories, and that learning theories are a diversion from ID. For example, Dick & Carey (1990) avoid mentioning any learning theory in their ID model. Wilson & Cole (1991) refer to this ID model that claims to be independent of learning theories '*procedural models for systems design*'. Wilson & Cole explain that *the procedural models often are represented as flowcharts reflecting a series of project phases, progressing from needs and problems analyses to product implementation and maintenance.* Branson & Grow (as cited in Wilson & Cole, 1991) argue that *procedural ID models depend less on learning theory and more on systems theory and project management methodologies.* Hannafin & Rieber (1989: 91) at that time saw little evidence of the relationship between learning processes and ECPs.

The second view, and the position adopted in this thesis, is that learning theories are implicit within ID even though the theories may not be mentioned explicitly. Wilson & Cole (1991) refer to these as instructional-strategy models, claiming that they are in concert with *Gagné's conditions-of-learning paradigm.* Gagné's conditions of learning presents a hierarchy of learning outcomes, each of which requires specific conditions for learning to happen. Therefore, the instructional designer has to know the desired outcomes and then identify and use appropriate strategies for learners to achieve those outcomes.

Hannafin & Rieber (1989) state that there are relationships between psychology and instructional design, while Thompson, Simonson & Hargrave (1992) show the impact learning theories has made upon instructional design. Hannafin *et al.* (1996: 379) add that computer use in education evolved interactively with developments in psychology. An example is the Collins-Brown cognitive apprenticeship model that is tightly linked to cognitivism (Wilson, 1995b). Wilson suggests that we should all look for praxis (interface between theory and practice), with psychologists providing knowledge on how learning could happen through technology, and designers looking at the best ways of instruction. It is important to note that it is possible that this view has gained currency with time. For example, it appears that Wilson & Cole have themselves changed towards the argument that learning theories are implicit in ID (compare their views in 1991 with those in 1996).

Relationships between learning theories and ID

Taking the later view, ID has undergone changes that are in concert with shifts in theories of learning (Tinker & Papert, 1988: 4; Winn, 1993; Wilson & Cole, 1996; White & Purdom, 1996), in tandem with improvements in computer hardware and software (Wilson, 1995b). Wilson & Cole (1996) show three stages of ID development which directly correspond with the three major learning theories, behaviourism, cognitivism, and constructivism. However, Winn (1993) thinks that there are four generations of ID evolution. These generations are important for one to understand the different models of ID, how each model might be used in class, bearing in mind that all models might be represented in a single programme, since, as I suggested earlier, a combination of learning theories might be implied in a single programme.

a. Behavioural models

Procedural or programmed instruction based on behavioural learning theories, for example of Skinner, Gagne and Rowntree, shaped the first generation ID, roughly during 1960-1975 (Hannafin, Hannafin, Hooper, Rieber, & Kini, 1996: 379; Wilson & Cole, 1996; Mergel, 1998; Jacobs, 1992: 117-118). Gagne (1985), Hannafin & Rieber (1989: 92-94), Reeves (1994), as well as Burton *et al.* (2001) articulate some of the ID behaviourist models. The expectations of behavioural ID models are behaviours learners demonstrate after using the programme (Hannafin & Rieber, 1989: 93). Thus, the learner's behaviour is predictable, can be conditioned through stimulus-response associations, using small units of knowledge and skills (Hannafin *et al.*, 1996: 379). These aggregate into a desired whole behaviour through reward and reinforcement without further intervention from designers or teachers. For example,

> *the designer writes behaviourally specific learning objectives, classifies those objectives according to a taxonomy of learning types, then arranges the instructional conditions to fit the current instructional prescriptions. In this way, designers can design instruction to successfully teach a rule, a psychomotor skill, an attitude, or piece of verbal information* (Wilson & Cole, 1991).

Learners start from easier to skills or concepts that are more difficult or complex (Wilson & Cole, 1991). Tinker & Papert (1988: 5) claim that such programmes are relatively easy to create, and easy to integrate into curriculum. Examples of this include simulated actions used to train aircraft pilots as well as simulated science experiments (Linn, 1988: 122-123).

The sole use of programmed learning died in early 1960s according to Mergel (1998) but in mid 1970s according to Wilson & Cole (1996), because it did not appear to live up to its original claims. Among others, Hannafin & Rieber (1989), Rieber (1992), Winn (1993), Schuman (1996), Alexander (1997), Greening (1998: 29), as well as Rieber *et al.* (1998), and Tennyson & Rasch (1988) summarise the weaknesses of behaviourist 'instructivist' pedagogy. They point out objections to, more importantly, the assumption that the teacher can see further than and for the learner; that it is only effective for low-level learning such as rote recall; and that each step presented is the best one to take in order for every user. These programmes take rules, definitions, and procedures as very important.

Regardless, Tinker & Papert (1988: 5) argue that behavioural approaches such as simulations, tutorials, drill-and-practice have are still useful in ID. Hannafin & Rieber (1989: 94) concluded that behavioural models are efficient. Nonetheless, the theorists cited in the above paragraph recommend a shift from behavioural to include cognitive and constructivist approaches.

b. Cognitive models

Hannafin *et al.* (1996: 379) and Wilson & Cole (1996) state that cognitive perspectives gained increased acceptance during 1960s because of the desire to inculcate cognitive processes, including assisting learners to form new concepts. Another essential for learning process is problem-solving (Wilson & Cole, 1991). However, according to Hannafin *et al.* (2004: 6-7), there are some similarities between behavioural and cognitive programmes. For example, content is broken down and ordered in hierarchy to meet externally determined objective, and knowledge and skills are conveyed through structured means.

However, unlike behaviourist models, the focus is on the individual, and how that individual selects, perceives, processes, and learns information (Hannafin & Rieber, 1989: 94). The model stimulates cognitive processing instead of teaching. For example, Lepper & Chabay (as cited in Wilson & Cole, 1991) note that some cognitive models accentuate learner-initiated inquiry, exploration, cooperative learning, and empathy, which traditional behavioural ID models do not emphasise.

Cognitive-based ID aims at learning that occurs as individuals construct 'schemata' that represents the world for them, and incorporates the notion of accommodation and assimilation (Gardner, 1983, 1993), and at matching learning to the individual's needs and style of learning (Cronbach & Snow, 1977; Tobias, 1976, 1989). Thus, the second generation is based on 'exogenous constructivism', by which the programme just helps learners with activities or exercises that makes them cognitively active towards new concepts and to better capacities to solve problems (Dalgarno, 2001: 185).

Features of cognitive ID models

This section outlines applications of cognitivism in ID that are relevant to this research. I draw notes from Hannafin & Rieber (1989: 97-98), Wilson & Cole (1991), Duffy & Cunningham (2001: 184), and Wilson, Jonassen & Cole (1993) who explain a 'cognitive apprenticeship' model. However, I consider Duffy & Cunningham's (2001: 184) warning legitimate, that the learner assumes responsibility although s/he is assumed to be an apprentice.

The cognitive model deals with content (as in textbooks) as domain knowledge (conceptual, factual, and procedural), but considers it insufficient to enable learners to approach and solve problems independently. The model also provides heuristic strategies that help narrow solution paths, for example, through repeated problem solving practice. However, the learner controls most of the activities. The cognitive model recommends situated learning: that is, learning that reflects the way the knowledge will be useful in real life or authentic contexts. Collins (as cited in Wilson & Cole, 1991, and in Wilson *et al.* 1993), as well as Duffy & Cunningham (2001: 179) advise that situated cognition should be based upon problem-solving situations. Collins gives an example in mathematics where learning could encompass shopping in a grocery store. A computer can be used to model such a situation, and the learners would be asked to articulate reasons for phenomena in the model. A teacher or an intelligent tutoring system gives hints to help (i.e., coaches) the learner when they are failing to solve the problem or if they are getting off-course. Cognitive ID strategies offer possibilities for transforming a learner's conceptual understanding, in a similar way practical work might.

Another important aspect of cognitive ID is that it can incorporate exploration, which encourages learners to try out different strategies and hypotheses and to observe the effects their trials. Collins (as cited in Wilson & Cole, 1991) claims that through exploration, students learn how to set achievable goals and to manage the pursuit of those goals - they learn to set and try out hypotheses, and to seek knowledge independently. Real-world exploration is always an attractive option; however, constraints of cost, time, and safety sometimes prohibit instruction in realistic settings. Computers offer additional advantages such as the ability to change the complexity or diversity of a situation instantly – this enables further challenges and offers grounds for testing concepts.

From about 1989 (Wilson & Cole, 1996), third generation ID started and incorporates the learner's inputs and control of the direction of learning. It is a generation in which the constructivism aspect in the cognitive theories ("Cognitive Complexity Theory", and the "Anchored Instruction" theory) as well as the "Instructional Transaction Theory" advocating for interaction (transaction) between learner and program are applied for discovery and experiential learning in computer "micro worlds"(Rieber, 1992; Wenger, 1987; Merrill, 1991, 1993). Learners chose what to learn.

The advancement in computers caters for Multiple Intelligence (MI) in third generation ID, providing instructional designers with many approaches to a topic. Examples from Tinker & Popert (1988: 6), Osberg (1997), and McKenzie (2001) include:

- **Linguistic** and communication tools: E.g., Word processing

- **Logical-Mathematical** (Logic and critical thinking skills): E.g., exploring, organizing data, programming, selecting relevant information and problem solving while playing a game
- **Visual/Spatial** and theory-building tools: E.g., microworlds, graphing utilities, modelling environments, creativity and visual skills; browsing through a 3-D programme
- **Musical:** E.g., composing music
- **Bodily-Kinesthetic:** E.g., hand-eye coordination with the keyboard and the mouse; moving objects around the screen
- **Interpersonal:** Working in groups in microworlds
- **Intrapersonal:** E.g., working independently at own pace

LOGO could be an example of a third generation programme; it provides different forms of activities in a non-linear format, creates exploratory environments, providing guided discovery and choices to the learner at any time (Linn, 1988: 127). Hannafin & Sullivan (1995: 19) are particularly impressed by the number of examples, amount of practice, review, and feedback in LOGO.

It is not clear where the third generation ends (and whether that is important), but beyond the third generation, use of computers in learning focused attention on interactive multimedia (BECTA, 2001), in which learners control what they do in a constructivist framework (Alexander, 1997), with the assumption that learners know best their needs. There seems to be cognitive constructivist models, which Papert (1993) argues are "dirty" (holistic and authentic), as opposed to behavioural approaches, which Papert terms "clean" teaching (isolate and break down knowledge to be learned). Cognitive-constructivist models view truth and knowing as local events, and highlight the importance of context and multiple perspectives in making meaning (Willis, 2000: 5), all of which can be disorganised (dirty).

Fourth generation ID and microworlds

Introduction

The fourth generation rejects cognitive science as the **only** (my emphasis) basis for instructional design, and the exclusion of the learner from planning or designing the learning experience. It relies on 'endogenous constructivism' by which learners discover and explore virtual environments (Dalgarno, 2001: 186). Constructivist experiences help learners to understand what they are studying (Salviati as cited in

Cunningham, 1991: 130), because, through participation, such experiences embody iterative use of knowledge and skills for further experiments and experiences (Winn, 1997). The design permits learners any kind of interaction the system is capable of (Jacobs, 1992: 119; Merrill, 1993; Young, 1996: 18), instead of prescriptions. The importance of context, and of social construction imply that **any** (my emphasis) model made by learners is just one of the many possible constructivist ID models (Willis, 2000: 9-12; Kozma, 2000: 13).

The fourth generation ID is specifically important as it ushers in the use of microworlds and open environments into ID, and offers opportunities for a wider range of learning strategies including constructivism. I give more detail in this section because Zadarh fits into the fourth generation ID, and is composed of virtual microworlds.

Microworlds, virtual environments, and virtual realities

Microworld

Jonassen *et al.* (2003: 90) state that Papert and the MIT Media Lab started the use of the term 'microworld'. A microworld is an exploratory learning environment that simulates phenomena, thus offering opportunities to learners to manipulate, explore, and experiment. A microworld is also known as a simple domain, focussing on the quality of a few interrelated constructs (Hannafin *et al.*, 1996:393). Microworlds are now presented in ECPs such as Zadarh. Other examples of microworlds or "phenomenaria areas" appear in Perkins (1991: 19), and include "aquariums", "SimCity", and "physics microworlds". Computer microworlds offer virtual environments and realities in which one can do many things, some of which are beyond reach in real worlds.

Wilson *et al.* (1993) explain that manipulating the equipment, the task, and the environment control the complexity in a microworld. Thus, *successful microworlds rely on learners regulating and controlling their own learning* (Jonassen *et al.*, 2003: 191). For example, microworlds can incorporate cognitive apprenticeships, which provide *opportunities for modelling, reflection, exploration,* and for a learner to reflect on his/her knowledge (Wilson *et al.*, 1993). Or can *contain adventure games, where players master each environment before moving on to more complex environments* (Jonassen *et al.*, 2003: 191).

The importance of microworlds is that they are more open for learning than, for example, laboratories, that are defined by pre-specified objectives. With these features, microworlds can qualitatively alter a learner's conceptions (Hannafin *et al.*, 1996: 393). Playing games in such microworlds is an example of applications of radical constructivism (Rieber, 1992: 94; Rieber *et al.*, 1998).

Before pursuing the debate on microworld, virtual environments, and virtual realities, I wish to draw attention to the terms virtual environment and virtual reality because they can be confusing. Indeed McLellan (2001: 457) indicates that the two terms are interchangeably used, together with cyberspace. McLellan notes that 'virtual' *denotes the computer-generated counterpart of a physical object*, and Rieber (1992: 94) describes VEs as ... *computer-based learning environments* ... Jonassen *et al.* (2003: 201) describe VR as ... *a type of microworld that provides learners with an interactive 3-D experience by surrounding them with a moving simulated world.* However, in this thesis, I take virtual environment as the space in which virtual realities happen or exist. A microworld generated by a computer then has space, which I refer to as the virtual environment (VE), and objects as well as activities, which I refer to as virtual realities (VR).

Virtual Environments (VEs)

Learners enter into an artificial microworld, which has VEs (Hannafin & Sullivan, 1995; Winn, 1996). The advantage of VE microworlds is that they cannot be provided by any other means (Winn, 1993). For example, computers can enrich VEs and extend our perceptual, tactile, and visual insight into concepts (Kiboss, 1998: 12). VEs can be used to teach science concepts, which are difficult to teach in real laboratories because learners interact iteratively with virtual objects in conditions, which are possible only in a virtual laboratory (Perkins, 1991; Ramsey, 1975:98-99; Dede, 1995; Winn, 1997; Geelan, 2000). Overall, access to knowledge and interactions are unrestricted in VEs, and offer open environments, possibly as described by Doll (1989: 246), that are useful for modelling (Stratford, 1997: 4-12), and encourage what Yore (2001) refers to as interactive constructivism. However, there are complaints that VEs have suffered prevalence of technology and aesthetics rather than promoting knowledge – they simply supply information without knowledge-building processes (Barbera, 2004, 14).

Open learning environments (OLEs)

I discuss open environments under ID because, as it will be seen, they are not often available in typical classrooms. OLEs are enabled in VE.

It seems that the notion of an open learning environment, as described by Doll (1989: 246), Hannafin, Hall, Land, & Hill (1994: 48), as well as Hannafin *et al.* (2004: 7), is constructivist since such an environment grants learners their wishes, but allows inputs from a facilitator and/or the programme. The constructivist design anchors learning activities to the learner's long-term or larger problems, but in a form authentic, and therefore open, to a learner (Savery & Duffy, 1995: 32-33). Rieber (1992:94), Hannafin (1999) as well as Savery & Duffy (1995) explain further that an environment is open if it allows a learner to chose interactions, goals, and /or the way to pursue those goals. The focus is on an individual's understanding, needs, perceptions, and experiences. Thus, Hannafin (1999) adds that OLEs guide learners to recognise or generate problems that relate to their needs.

Constructivist OLEs are chaotic

Openness brings with it multiple demands, since each learner might have different desires and methods of learning. Thus, Wilson (1996) believes that among the difficulties with open environments is the possibility that they might be fuzzy and ill defined, but argues that an environment that is good for learning cannot be fully packaged and defined. Learners might chose activities, pace and direction, to the extent that the end outcome is uncertain and uncontrolled. Thus, Winn (1997) as well as Hannafin (1999) point out that strategies for providing guidance, feedback to actions and collaboration, are not so straightforward.

Wilson (1996), Dede (1995), and Perkins (1996) note differences in the amount of guidance or direct instruction found in learning environments, and observe that varying degrees of guidance pose different instructional challenges. According to Wilson, the teacher or instructional designer has to be tentative to accommodate learner freedom. Learners can be provided with perspective-setting or -altering contexts that help to activate relevant prior knowledge, experience, and skill related to the problem and to potential strategies to be deployed (Hannafin, 1999).

Unfortunately for an evaluator, Wilson observes that the same chaos (desirable in OLEs) is also characteristic of poorly designed OLEs – i.e., it might be difficult to know whether the chaos is intended or is a result of poor design. For example, when learners get lost or stranded, one need to find out whether this is designed to help them solve a problem or it is due to lack of support.

Virtual reality

The advent of virtual reality (VR) boosted the fourth generation programmes (Winn, 1993). Computer-generated microworlds provide VR in which there are opportunities for exploration (Cohen, Tsai, & Chechile, 1995) of phenomena that would be difficult, or intangible under usual laboratories. Additionally, traditional lessons sometimes lack real-life analogies on which to build mental models, because there are no such events in the real world (Dede *et al.*, 1997).

Jonassen *et al.* (2003: 201) explain that an outstanding feature of a good VR is 'immersion'. Dede (1995), (Osberg, 1997), as well as Moshell, Hughes, & Loftin (1999) add that immersion in VR can provide the subjective impression that one is participating in a "world" comprehensive and realistic enough to induce suspension of disbelief. That is, the user becomes isolated from the real environment and interprets the images in the VR as being real. This makes the user interact intuitively like an inhabitant of the VR.

According to Dede (1995), Zeltzer (1992), and Dede, *et al.* (1997), VR improves learners' understanding relative to other technologies because VR accommodates autonomy, presence, and interaction. That is, VR can engage learners with experiences, which facilitate perceptual experiences. Thus, VR supports constructivist learning (Greening, 1998). The theorists in this paragraph also believe that another useful characteristics of VR for learning is that it motivates a learner by inducing him/her to spend more time and to concentrate on a task. I think these are ways the designers of Zadarh used VR.

VRs face difficulties of cognitive load. For example, there is a difficulty of *switching attention between the different senses for various tasks* (Dede, 1995). Hence, Dede (1995) advises for taking care of speed.

Games in VE and VR microworlds

Various authors (E.g., Linn, 1988: 128; Leutner, 1993: 113; Tinker & Papert, 1988: 3-7; Greening, 1998; Dede, *et al.*, 1997; Rieber, 1996a) elaborate on constructivist microworlds that include games. The process of playing the game is constructivist in that the learners are co-designers (design to learn kind of approach), but should be guided. These authors advise that the game should be interesting with graphics that is appealing and music. There has to be a defined start and finish of the game, and an overall floor plan or map.

Constructivist ID as a change agent for the teacher-learner relationships

I liken the open and constructivist virtual environment to a nature reserve where one is free to explore, only that the virtual environment is safer. The learner has a platform to challenge held constructs of his/her own and those held by others, including the teacher. Jonassen *et al.* (2003: 13) points out that the *teachers must relinquish at least some of their authority, especially their intellectual authority.* This is because *students should construct their own meaning…* I wish to add that, in the context of a computer as a new tool in disadvantaged communities, the learner might explore it more than the teacher, thus levelling the skills authority. Therefore, a constructivist open virtual environment or microworld, as provided in Zadarh can be a starting point for changing the teacher-learner intellectual and skills relationships. This in turn could act as a change agent for the teacher.

Instruction and construction?

It is notable that instruction and construction appear to be antagonistic such that there could be tension between classical (traditional) ID and the newer radical constructivist approaches in ID.

Hannafin *et al.* (1996: 395) as well as Hannafin, Hannafin, Land, & Oliver (1997: 107) discuss this apparent tension at length. For example, Hannafin *et al.* (1997) argue that Gagne's instruction model takes *reality as objective and independent of the individual learner.* On the other hand, constructional design creates environments in which a learner can design his or her own tasks and constructs. Thus, *the term 'instruction' is considered a pejorative to some in describing emerging learning systems* (Hannafin *et al.*, 1996: 395).

Hannafin *et al* (1997: 113-114) and Schuman (1996) advise inclusion of aspects of each learning theory because each theory has strengths and weaknesses. In support, Ertmer & Newby (1993) and Davidson (1998) point out that learning theories are compatible with ID, such that Scott, *et al.* (1987) talk of the application of constructivism in instructional design with reference to games.

Application of ID in assessment - Computer-Aided Assessment (CAA)

Introduction

CAA offers ways which provide opportunities for individual self-assessment any number of times, and provides data immediately in formats that teachers can use to remedy learners' problems, however large the number of learners are (Bugbee & Bernt, 1990: 98; Oliver, 2000: 1; Thelwall, 2000: 40, 45, 46; Gretes & Green, 2000: 46; Croft *et al.*, 2001: 62). Additionally, CAA programmes such as QM can calculate and interpret learners' performance, and ensure reliability because computers mark the same way all the time. Thus, technology *promises to improve assessment practice* (Hickey *et al.*, 2003: 531).

However, studies concerning CAA use are inconclusive, and many questions about its implementation in schools require more research (Gretes & Green, 2000: 47; Thelwall, 2000: 46-47). Schools in SA that I am acquainted with have not used and evaluated CAA. A study such as this one is necessary to examine the potential of CAA and the way CAA can address some of the problems that militate against diagnostic assessment, especially of a large number of learners in a more rigorous manner.

Scenarios of CAA/CAT use

At a school, CAA can be made available on a Local Area Network (LAN) or on-line (in which tests are done and marked on a web site). Oliver (2000: 1) gives a *CAA* scenario in which learners 'log-on' to a LAN using a password, and CAA indicates a test together with the number of questions, and provides information on the mode of answering those questions. Proper CAA, according to Oliver, stores the learners' answers, mark, and keep the results of the test accessible to those who need it, in a secure file format. Many CAA packages include statistical formats, within a secure 'staff only' file, and the capability to direct the learners towards further work, either remedial or complex, on the basis of their performance. The CAA could be summative or formative CAA, including grading tests, open access test, self-test, exercises, and diagnostic tests (Thelwall, 2000: 39). Many theorists (E.g., Olsen, Maynes, Slawson & Ho, 1989: 312; Wise & Plake, 1990: 4; Dekker & van Niekerk, 1991; Sandals, 1992: 71; Twomey & Miller, 1996: 5; Tamir, 1996: 123; Bennett, 1998; Gretes & Green, 2000; Gathy *et al.*, 1991: 109; Olsen, 1990: 36) give wide-ranging roles that CAA can play, which paper-administered tests cannot accomplish. These include own-pace self-assessment of learners, and immediate feedback, thus making diagnostic assessment and remediation of learner problems possible. Other roles of CAA might improving test validity, storing large item banks from which tests could be set randomly, and easing statistical analysis (for example, the difficulty and discrimination indices, and reliability of the test). Computers do not get tired as do teachers, so continuous assessment is easier. Computers also keep learners' marks confidentially and so reduce the embarrassment learners may feel under the more traditional classroom environment if they perform badly.

Possible roles of CAA in science classrooms

CAA carries with it the advantages mentioned in the above. However, CAA might be incorporated in constructivist environments, where learners can be tested on practical skills.

Possible disadvantages of CAA in the contexts researched

Bugbee & Bernt (1990: 97), Bernt, Bugbee, & Arceo (1990: 271), and Wise & Plake (1990: 7-9) report that computer technology might negatively affect learners and that there is evidence that the learners could achieve lower marks if they use CAA for the same test items. The reasons for these include computer anxiety or attitudes learners have towards computers, screen glare, lack of computer skills, and the difficulty learners could have in reading text on the screen of a computer monitor. Wise & Plake report that learners take a shorter time using a computer to do test, and that providing immediate feedback increased anxiety and lowered test scores.

Additionally, Taiwo, (1995: 19), Bugbee & Bernt (1990: 92), Sandals (1992: 68, 70-74), and Thelwall, (2000: 40), mention problems that might be prevalent in South African disadvantaged communities. These problems include the high cost of computers and of CAA software, system speed and breakdown, the possibility that it might not be adequate to assess higher order thinking skills, and lack of computer skills among teachers. There is also the fear that learners could give their passwords to others to do the test for them or learners glancing at other screens where computers are so close. I have noted as a student at university, and at my work that the performance of computers is also related to the expertise available to set up the computers properly, and to help those learners who experience problems.

Therefore, Oliver (2000: 2) advises against the unnecessary use of CAA where other simpler means of testing would be sufficient. For example, CAA might take more time to set up, and administer, although the teacher might recover some of that time from the efficiency and speed of CAA especially for large numbers of learners (Thelwall, 2000: 41; Olsen *et al.*, 1989: 313).

Using computers to diagnose

While CAA is at the forefront among alternative assessment tools (Kumar, 1994: 59), the potential of a computer to diagnose has rarely been taken advantage of (Bright, 1987: 85). Bright (1987: 75) explains that using computers to diagnose is *most appropriate* when a learner *has repeatedly been unable to learn,* and so, there should be suspicion of *fundamental misunderstanding* that needs to be corrected, not forgetting that a learner may have a learning disability. Bright (1987: 77) notes that while teachers would evaluate unexpected learner responses against earlier responses, using computers to evaluate is limited by the capability of the program. One of the important considerations is that a diagnostic CAA program has the capacity to provide remedial responses that help and make sense to the learner. On addition, the ability for CAA to diagnose, as with other forms of assessment, is increased` by instantaneous marking. Bright (1987: 81-83) advises that even then, teachers ought to do further remediation off-line. Bright (1987: 77-79) provides further advise on how teachers could work with diagnostic CAA programs:

- *The CAA diagnostic program could match particular responses with errors pre-identified and stored in the program;*
- *A procedure must be developed for turning a learner's wrong responses into a specification of the errors being committed;*
- *The follow-up on remediation is enhanced by a CAA program, which can keep records of the learner's performance and the particular diagnosis that has been done previously; and*

- *A change in a curriculum, might change the nature of mistakes and errors. Therefore, there is a need for vigilance and continuous re-configuration of diagnostic procedures.*

CAA software

CAA software is currently big business and is of intense academic interest that many are advertised on the Internet. The Internet also provides sites such as Software Reviews, which provide reviews of CAA software. Institutions also have reports regarding the use of various CAA software, while authors such as Booker (1998), Oliver (2000: 2), and Twomey & Miller (1996: 7) provide links to CAA software packages available, some commercial, others free. Examples include "EQL Interactive Assessor" (commercial), and "Markin32", (free), for marking essays online.

The authoring programme that I used in this evaluation was QM. Question Mark has been used extensively in CAA (for example, Knight & Brown, 2000; Thelwall, 2000; Croft *et al.*, 2001). QM has two important facilities. The one is Question mark Designer (QM Designer), which provides shells into which one can type MCQs. QM Designer offers options of allowing learners to repeat a test, showing learners their marks, and showing learners the time they have spent doing a test. The other facility that is important for diagnosing teachers' and learners' problems is Question Mark Reporter (QM Reporter). QM Reporter analyses and makes various reports, such as individual performances and number of trials, class average, test performance, etc.

Previous use has shown some difficulties (Knight & Brown, 2000: 2, 5; Thelwall, 2000), but these might be related to the version used, the experience of the users, and the nature of computer hardware and software. For example, it was found that QM could not import documents from word processors such as 'word', as well as diagrams and equations, and produced poor graphics. Despite these problems, Knight & Brown (2000), Thelwall (2000), and Croft *et al.*, (2001) found that QM was able to deliver CAA with most of the advantages of CAA mentioned in the above.

Examples of intercourse between science and ID

Introduction

The first intercourse is that the designing of instruction might share some processes, such as problem solving. The other intercourse, and the one I deal with in this section, is about programmes that instructional designers have made in science. Many of these model or simulate science concepts or processes.

Simulating models for science concepts

Simulations are useful in confronting students' misconceptions to promote conceptual change such that simulations can be integrated into inquiry activities (Rieber, 1996b: 16; Hendricks, 2002). Teachers use simulations in science experiments. However, computer technology has increased the ability to design simulations that are more sophisticated. An example is the Genetic Construction Kit (Thomson & Stewart, 2003).

Dalgarno (2001: 186) claims that there is no accepted definition that would distinguish simulations from microworlds. This claim might be because most simulations happen in some form of microworld. Nonetheless, a *simulation focuses attention to the reality it purports to model* (Stratford (1997: 4). *Simulations are a powerful form of Computer-Assisted Learning (CAL) as they have the potential to enable the user to carry out experiments that would otherwise be impossible* (Watson, 2001: 587), and a *number of studies have emphasised that software ca be used to make modelling 'accessible' to students in classrooms* (Rieber, 1996b: 16). Stratford reports successes from using simulated models of the Mendelian genetics, while Rieber (1996b: 13-16) cites where animated graphics were more beneficial than text.

However, Rieber (1996b: 17) believes that large-scale integration of such programmes into school systems is difficult and concludes that there is no firm consensus on the benefits to learners from running simulations or constructing models. According to Rieber, part of the problem is the relative scarcity of research into computer-based models.

> *This in turn is due to the relatively recent introduction of the computer into classroom use, the lack of easy-to-use models and modelling environments, the difficulty of collecting and analysing data generated by learners engaged in using computers, and the huge investment of time and effort involved in designing and implementing model-based software research (Rieber, 1996b:17).*

Hence, Rieber suggests to evaluators to check on, *inter alia*, whether programmes and models in them are oversimplifying reality, what prior knowledge learners need before using a programme, what relationships are created between mental models and the models in the programmes, and what confidence learners put in the validity and reliability of models they construct. Furthermore, Rieber (1996b) recommends studies to find out how learning environments and modelling might accommodate diverse learning styles and preferences.

Implications of learning theory-ID relationship for evaluation

Evaluating a programme against performance requires consideration of the learning strategy upon which the programme is based. Hickey & Zuicker (2002) argue that the evaluation of an ECP is embedded in the views of knowing and learning of the stakeholders.

That is, a behaviourist ECP would probably be evaluated with empirical evidence (Hannafin & Hannafin, 1991:303), while a constructivist programme would take cognisance of the views of participants and how they go about developing those views. Alternatively, I suggest that an ECP could be evaluated against the different learning strategies, since it might support all of them.

Wilson *et al.* (1993) show that a cognitive-designed programme should check upon thinking processes as learners interact with the programme. Such processes can be checked by think-aloud protocols (Smith & Wedman, as cited in Wilson *et al.*, 1993; Dede, 1995; Middleton, 1992: 254). The learners' verbal reports become a source of data for making inferences about their actual thinking processes, which in turn provide evidence concerning the effectiveness of the programme. Wilson *et al.* also suggest that learners might also be asked to elaborate on their verbal reports, particularly the reasons for decisions, which result in errors.

A seemingly, constructivist-based approach is outlined in Dede (1995) and in Middleton (1992: 254). They suggest that evaluation should be continuous to check on *usability, motivation, and the learning process*. By the use of talk-aloud protocols, a cycle of prediction-observation-comparison can be used to monitor the learning process, as well as for identifying usability problems. These procedures should be accompanied by a careful initial analysis of learner needs and capabilities/limits of the technology.

Another pedagogical concern is context. Wilson, *et al.* (1993) argue that success should be measured in the context the programme will be used. This involves environment analysis, by which the context of the instructional system, both where the instruction will occur and how the instructional materials will be used, is investigated.

Hannafin & Rieber (1989:98) offer adequate advice in stating that *the issue is not, which models are best, but which design decisions are most appropriate given the demands of the learning task.* Each model has the potential to achieve some form of learning, and what matters is what form of learning a programme is evaluated against. Gredler (2001: 521) warns that a programme could be wrongly evaluated or compared to another method of teaching. Gredler gives an example of comparing simulations against the lecture method, and states that the two achieve different kinds of learning.

With particular focus on computer games, Gredler (2001: 537) advises evaluators to look at the design validity (E.g., knowledge domain and subject area expertise), cognitive strategy, and/or social interaction processes executed by learners during play, and following-up on specific processes and effects. Therefore, evaluation of ECPs has to look at the appropriateness of the programme to the intended outcomes, bearing in mind the claims designers make.

Considerations for evaluating ECPs

Generally, in principle, evaluation is important for making decisions on curricula issues, which include materials, objectives, contexts, and benefits such as learner achievement (Madaus & Kellaghan, 1992: 127-147; Munro, 1975: 220 – 221; Weiss, 1998). Many evaluations of ECPs are done against needs (Castellan, 1993: 233; 10; Fink, 1995: 10; The Maricopa Center for Learning & Instruction, 2000; Hitchcock & Hughes, 1995: 31; Percival & Ellington, 1984: 100) so that ECPs can be effective and consistent with solving the problems of the users (Richey & Nelson, 1996; Fink, 1995: 10; Imenda & Muyangwa, 1996: 35).

Evaluations provide important feedback to the designers of the programmes, bearing in mind that infusing technology in education is complex and not entirely captured by traditional causal relations because technology may be only one of many input variables causing changes, and many variables cannot be controlled (Bodilly & Mitchell, 1997: 16-20).

It is recommended that evaluation in ID should be of a formative, cyclical, and interactive in nature. Each cycle provides data and feedback from the experiences to new cycles, showing how the different components of a programme are working, and leading to improvements in efficiency and quality (Fink, 1995: 10; Fuchs, 1995: 1; The Maricopa Center for Learning & Instruction, 2000). Formative evaluation is used to increase the efficiency and effectiveness of their instructional materials (Dick & Carey, 1978:158). It could be in that light that Nova (1986: 17) identifies formative evaluation among the essential activities in education. Formative evaluation can involve expert review (E.g., scientists), one-on-one or small group trials, and tryouts with the target audience (E.g., learners) under the conditions that the materials were designed to function in (Wilson *et al.*, 1993).

Among the methods that are recommended for formative evaluation, is the systemic approach, which focuses on identifying working and ineffective parts of the programme. A typical systems evaluation appears in Percival & Ellington (1984: 114), called 'error elimination' advocated by one Karl Popper. Error elimination is both open ended and ongoing, and involves testing the instructional system for errors, and identifying the new problem situation. Castellan (1993: 234–235) among others, includes these aspects for evaluation:

- *Technical accuracy* – whether software executes correctly and accurately on learners' platforms
- *Substantive fidelity* – concerns accuracy of content and whether the content is worth learning.
- *Integrative flexibility* – this questions the compatibility of the program into the curriculum
- *Cyclic improvement* – technology should be continuously improved

Besides the evaluation of systems compatibility, a programme can be evaluated for its value to the potential end-users, and for outcomes, they achieve when they use the programme. Some outcomes are immediate; others are long-term (Bodilly & Mitchell, 1997: 20; Nevo, 1986; Madaus & Kellaghan, 1992). In order to get at all these values and outcomes, multiple measures (quantitative and qualitative) should be used. According to Heinecke *et al.* (1999: 1-2) these might include (I selected the questions which I think are particularly important to this research):

- **Valuing**: Is this a good program? By whose criteria of good should we judge a social program? ... Should programs be compared to each other or to absolute standards of performance? Should results be synthesized into a single value judgment?
- **Knowledge construction**: How do I know all this? How complex and knowable is the social world? What are the consequences of oversimplifying this complexity? Does any epistemological or ontological paradigm deserve widespread support? What methods should evaluators use?
- **Evaluation practice**: Given limited skills, time, and resources, and given many possibilities, how can I narrow my options to do a feasible evaluation? What questions should I ask? What is the role of the evaluator? Whose values should be represented in the evaluation? Which methods should be used? What should the evaluator do to facilitate use? What guides these choices?

These concerns are reiterated in Hickey & Zuicker (2002: 541). For example, they mention the limitations of time and resources, the need for comprehensive evidence given limited access to classrooms, and the difficulty in drawing conclusions from diverse methods.

Many of the values and outcomes relate to the performance of learners and issues of the curriculum. One of the important aspects of a curriculum is evaluating pedagogical soundness (Castellan, 1993: 234–235). Concerns include, *inter alia*, articulation of instructional goals, appropriateness of technology to the concepts learned, timing of application, and whether the programme encourages exploration, testing, application of ideas and concepts, and self-assessment.

Audiences and participants in an evaluation

Teachers are responsible for judging the progress learners make. Therefore, *evaluating curriculum programmes has been seen as one of the main areas where* *teachers can undertake school-based inquiry* (Hitchcock & Hughes, 1995: 32), presumably because teachers are likely to be the implementers of evaluation findings. In this research, the targeted beneficiaries of the evaluation of the two computer programmes were the teachers and learners; therefore, they were the ones to decide whether these programmes were good, and so, they had to participate in planning some aspects of this research and in setting criteria of merit for judging the value of the two programmes.

Hannafin, Reeves, & Hayden (2001: 252) believe that evaluators tend to focus on local or specific learners, classes or schools. In this way, evaluators sometimes overlook the interests and views of policy-makers. Additionally, interests in a applying a programme in a school go beyond the school, and for these interests, Hannafin, *et al.* (2001: 252) argue, *evaluators have little control over either the goals or questions.* Key stakeholders determine goals and questions. *In all evaluation contexts there are multiple, often competing, potential audiences – groups and individuals who have vested interests in the programme being evaluated, called stakeholders…* (Greene, 1994: 531).

Audiences of evaluation include the researchers, participants in research, the community of scholars, and those who inherit the knowledge generated by the research (LeCompte, Preissle, with Tesch, 1993: 316), who might be interested in different questions and answers (Trotter, 1998), such that there could be little or no consensus on the evaluations. Hence, Castellan (1993: 236) states that there is no clear correct answer to research questions. For example, Hannafin *et al.* (2001: 253) argue that policy makers are interested in summative data, and such data, according to Trotter (1998), could be from a cost-benefit model in which scores (test results) are viewed as indicators of benefits. Thus, evaluation is not neutral (Hitchcock & Hughes, 1995: 31; Heinecke *et al.*, 1992; LeCompte, Preissle, with Tesch, 1993: 316), and one can find disagreements among evaluators [and also stakeholders] regarding the most appropriate approach because evaluation involves audiences with different interests (Madaus & Kellaghan, 1992: 132).

Thus, evaluators must negotiate on what is evaluated, whose interests or questions are addressed by an evaluation and against which standards the evaluation is made (Greene, 1994: 530-531). On page 531, Greene claims that *it is the fundamental* *political nature of programme evaluation contexts, intertwined with the predispositions and beliefs of the evaluator that shape the contours of evaluation methodologies and guide the selection of a specific evaluation approach for a given context.*

Conclusion – points to note in evaluating ECPs

It appears from literature above that the debate on the use of ECPs in science classrooms is continuing, especially because their advantages are not clear. However, more programmes are being produced. At the beginning of this thesis, I took the position that the ECPs are wrongly evaluated – they are evaluated against outcomes for which they are not designed, and that the evaluations are done for the end-users, instead of being done by the end-users. These concerns are disturbing in communities that are disadvantaged, who nonetheless wish to get on the information technology trek. Costs of information technology has excluded these communities, and, as is the case with many well-equipped schools, have had their teachers and learners excluded from judging the worth of ECPs. They are often observed by experts go to a school and observe as they try ECPs, and then conclusions are drawn for them. This evaluation aimed at end-users owning the evaluations.

In the next part of this thesis, I give the philosophies, methodology, and methods, which I think could be all-inclusive in terms of using comprehensive methods of evaluation and in involving teachers and learners in the evaluations.

PART IV

EVALUATING COMPUTER EDUCATIONAL PROGRAMMES FOR CHANGE

The process of evaluating computer-aided educational programmes in South African disadvantaged communities involves guiding teachers towards identifying their problems in the classrooms, prioritising their needs, and enabling them to evaluate programmes. This can be easier if teachers are well informed about the subject area a programme deals with, and about the way that subject is taught.

The facilitator of the evaluation process has to check on the status of computers in schools, train teachers where necessary, while simultaneously installing programmes.

Introduction

I wish to reiterate that one of the major rationales for this evaluation was to get teachers to participate in judging the worth and utility of two ECPs, with the hope that the experience would change the classroom practices of those teachers. I have also already proposed that previous evaluations have led to mixed findings about the success of ECPs, and to a misuse of ECPs in schools because other people carry out the evaluations on behalf of the teachers. Teachers and learners have previously been excluded audiences. The focus therefore, was on the participants, besides CAA and Zadarh.

This section explains the factors I considered important, and the philosophies, methodologies, as well as methods, which I used to see to it that teachers benefited from their participation in the evaluations.

The importance of this evaluation

The previous chapter argues that an evaluation serves different audiences differently. Besides, the general uses of evaluations outlined in the previous section, this evaluation aimed at equipping participants with skills of judging the worth of ECPs, as well as increasing their awareness of their classroom problems, and of the intricacies in integrating computer technology into the school curriculum.

Evaluation and research

There is debate on whether evaluations are research activities (E.g., Paton as cited in Hickey & Zuiker, 2002: 541). The similarity or differences are important in selecting the conceptual framework of an evaluation. For example, if research and evaluation are the same processes, then we can ground an evaluation in research philosophies and methodologies.

Different philosophies view research differently. For example, a seemingly positivist definition of research is *the systematic, controlled, empirical and critical investigation of hypothetical propositions about the presumed relations among natural phenomena* (Kerlinger cited in Cohen & Manion, 1987: 5). In social sciences, research is an investigation of phenomena surrounding human beings (Denzin, 1994; LeCompte, Preissle, with Tesch, 1993). Social scientists point out the difficulty in placing human beings under controlled experiments and in using numbers (empirical data) to describe human behaviour.

Generally literature describes evaluation as the systematic investigation of the worth or merit of some object, program, project or materials based upon a set standard and criterion (The Joint Committee on Standards for Educational Evaluation, 1994; Fink, 1995: 2). Such investigations involve gathering information (National Council of Teachers of Mathematics, 1995: 1; Hitchcock & Hughes, 1995: 31, Greene, 1994: 530-531) upon which decisions are made (Muraskin, 1993: 4). A course or material is evaluated by assessing the desirability of outcomes gained when it is used (Lloyd-Jones *et al.*, 1986: 1; Weiss, 1998: 5).

Some authors differentiate research from evaluation. For example, MacDonald (as cited in Elliot, 1991, 218) points out that research generates data for abstract theory while evaluation seeks the needs, and answers practical questions. However, from the above definitions, both research and evaluation involve systematic methodology, such that Hickey & Zuicker (2002: 541) define evaluation as *the application of research methods to inform a broad audience about programme effectiveness.* My view is that, as research does, evaluation can also generate data for theory on evaluation. It also appears that we do not hear of philosophies dedicated to evaluation only, and an attempt to evaluate forces one to adopt research philosophies. Thus, evaluation is in fact a research exercise (LeCompte *et al.*, 1993; Hitchcock & Hughes, 1995: 31), during which the worth of programmes can be systematically investigated although not all researches are evaluations. This is the view I take in this thesis, and so I use research philosophies, methodologies, and methods in evaluation. I explain the way I understand and applied the philosophies, methodology, and methods I used below.

Philosophical grounding of this evaluation

The requirement for evaluators to outline their theoretical and epistemological dispositions is frequently emphasised (Huberman & Miles, 1994: 428-429; Myers, 2000: 3; Pitman & Maxwell, 1992: 765). These authors argue that dispositions show how an evaluator construes the world and that dispositions influence the collection and interpretation of data. The outline of the philosophies will help readers understand why I designed the evaluation process the way it is. Furthermore, stating the dispositions contributes towards trustworthiness (Maykut & Morehouse, 1994:145-147), since they provide grounding for methodologies and methods.

What informed the choice of the philosophy and methodology used?
I have to point out that most of the teachers who took part in these exercises were also participating in teacher development programmes such as the Dinaledi project and Jula in SA, or were furthering their studies for Advanced Certificate in Education at the University of Natal or Rhodes University, in which I was one of the instructors. Therefore, I, to a considerable extent, knew the participants, their desires, qualifications, and problems and was in good talking terms with them. I am mindful of the possible bias due to using teachers who are already motivated towards further knowledge. Anyhow, I assume that an accurate evaluation would involve understanding the participants.

The fundamental benchmark was making sure that data represented as much as possible a true picture of what went on during the time participants used the programmes. This is in concert with Eisenhart & Howe's (1992: 657 – 662) and Heron's (1996: 159) advice that it is important to be clear about the grounds of validity and to be critical about the extent to which those validities have been reached. The rest of the aims of this evaluation would be realistic only if data was accurate. I considered the following measures basic grounds upon which a valid evaluation can be conducted.

a. Accessing and incorporating into the evaluation process the participants' values
The first consideration was whether the evaluation framework fitted *the research situation as viewed by the stakeholders* (LeCompte *et al.*, 1993: 322-349). I considered teachers and learners to be the most important stakeholders and participants, and therefore their values, which I think influence the way they view and judge a process, activity or product.

I adopted Lincoln & Guba's (1985: 160-161) definition of a value: *...criterion ... that one brings into play ... in making choices or designating preferences.* Lincoln & Guba explain that criteria include cultural or social norms, and these *are regulators of* *thoughts, feelings, and actions imposed by society or a cultural group on its members.* I had to choose philosophies that were as much as possible compatible with the value systems of the participants' cultures and of their schools.

An example is '*Ubuntu*' in those communities. I belong to the '*bantu*' group and grew up in that value system. Among other properties, '*ubuntu*' is a communal way of dealing with problems and of respecting people. That is, I am duty bound to see that I help people (and they in turn help me), and that we greet, communicate, and agree on how to sort problems out. This '*ubuntu*' is in agreement with some elements of philosophies in research that offer platforms for negotiated processes; where a researcher agrees with participants as equal partners on aims, solving problems, and representations of the meanings or realities.

The close relationships between me and participants, and agreed-upon nature of participation of consenting teachers and learners imparted into the evaluation process some '*ubuntu*', and some of the participants' interests, and made the evaluation more value-bound, ethical and authentic. In other words, the evaluation did not subscribe to positivism, which claims to be value-free. The NRF (2004) and Elliot (1991:217) support this approach, which at the same time satisfied 'Internal Value Constraints' (LeCompte *et al.*, 1993: 322-349; Erlandson as cited in Heron, 1996: 159).

b. Empowering teachers with skills in evaluation and computer use in science classrooms (value criteria)

Heinecke *et al.* (1999), LeCompte *et al.* (1993: 316-321), and Greene (1994: 533) argue that the questions and values that are addressed and promoted are important in determining the framework and methodology. Besides, exercises that benefit teachers are likely to be more successful (Martin, Hawkins, Gibbon & McCarthy, 1988: 185; James, 1988: 189). Similarly, the NRF (2004) considers teacher education and development as very pertinent, arguing that teachers play a pivotal role as pedagogical agents of change.

This consideration is linked to 'a' in the above, in that the participants' ought to value the skills and empowerment they anticipate to obtain from the study. Negotiations as indicated in 'a' revealed that participants were excited about improving their conceptual understanding of science, learning how to use ECPs, and gaining more skills in evaluation. One way of incorporating these values into the evaluation process was to design it in such way that teachers and learners drove the evaluation and reported on their experiences, and that my duty was to facilitate the evaluation process. In the long term, this would contribute towards personal and social transformation of participants (Heron, 1996: 170; Elliot, 1991: 231; McKenney & Van den Akker, 2002: Page III 407; LeCompte *et al.,* 1993: 326; Pachler & Byrom, 1999: 126).

Hence, the evaluation process had to incorporate training through workshops or dialogue, and had to grow partnerships between the participants and myself. Introducing the programmes, the training of teachers, and the evaluation process happened simultaneously. An open agenda whereby teachers could reach me, ask questions, and discuss the problems they were facing with the programmes helped me assess their knowledge on the use of computers in education, their knowledge about evaluation, and the knowledge of concepts in the programmes while at the same time helping me to refine the questions in the evaluation scheme. Training and providing information to participants during research or evaluation is implied in interpretive and constructivist evaluations (Heinecke *et al.*, 1999; LeCompte *et al.*, 1993: 316-321; Greene, 1994: 533, Hitchcock & Hughes, 1995: 31; Guba & Lincoln, 1989: 44; Kuiper, 1997; Kasalu, Doidge, & Sanders, 2002). In the explanations of Muraskin (1993: 5), this was both a process and outcome evaluation; teachers and I examined the evaluation process as well as the direct effects of the programmes on participants. It was an 'evaluation to learn' process.

c. A need for technical and instrument validity

The fit between research questions, data collection procedures, and analysis techniques or the suitability of data collection techniques or instruments with the type of data required and research questions formulated is known as technical or instrument validity (Eisenhart & Howe, 1992: 657 – 662; Hitchcock & Hughes, 1995: 105-106). In these evaluations, I had to make sure that the evaluation instruments and descriptions of findings were appropriate, and accurate, and that participants understood them and the assumptions imbedded in them (Sanders & Mokuku, 1994: 483; Eisenhart & Howe, 1992: 657 – 662; LeCompte *et al.*, 1992: 322-349;

Huysamen, 1994; Hitchcock & Hughes, 1995: 105-106; Heron, 1996: 159-170; Denscombe, 1998: 213-214; Reeves & Hedberg, 2003: 34).

These validities were related to the kind of data I was looking for. In view of 'b' in the above, I wanted data that would reveal the participants' level of competency, in identifying their problems, in evaluating ECPs, in conceptual understanding, and in how it can be introduced into the school curriculum, as frequently as possible. In the final analysis, I wanted data to reveal the values teachers and learners attached to the two computer programmes.

I reiterate that *inter alia,* training was meant to increase the participants' ability to judge the appropriateness of programmes and instruments. In addition, the participants and I had to go through the instruments and make sure that we had a common understanding of the instruments and of the roles of each one of us. This sometimes meant resorting to simpler English language or where necessary, to translating statements into vernacular. Triangulation, that is using different methods to extract data from the same experiences, improved these validities.

d. Consideration of the learning theories and strategies employed in designing the programmes

I feel that the evaluation is made more valid if the learning theories adopted in the programmes are in concert with the philosophy of the evaluation process, particularly considering the fact that evaluation is an integral part of Instructional Design. We have to be aware that embedded within the programmes is also the nature of the subject it is presenting, and therefore the philosophy and approaches used in that subject. For example, the constructivist programme Zadarh might require constructivist evaluation procedures. Secondly, the appropriateness of learning theories in the programme ought to be evaluated too against those the educational authorities recommend. This is why I outline the learning theories and the philosophies in science predominant in South African schools, as well as practices in South African disadvantaged classrooms in this thesis. These considerations undoubtedly complicated these evaluations.

e. The importance of participant's views

One consideration that influenced the choice of philosophies and affected the validity *of this evaluation is that I assumed that participants had views, which differed* between themselves, and between them and other stakeholders. I considered these views very important because participants would eventually be the users of the computer programmes. I am supported by LeCompte *et al.* (1993: 315-316) who argue that *... kinds and degrees of truth are held ... differentially for different audiences and constituencies.* That is, the epistemological position was that the participants' realities are paramount, and therefore, the ontological position was that knowledge is primarily subjective to the holder and secondarily objective only when the subjectivities lead to an agreed meaning.

The research philosophies chosen for this evaluation

The considerations above drive this evaluation with a social agenda and take the concerns of social development of teachers and learners much more seriously. The evaluation sought the participants' cooperation, realities, constructs and ideas, and so relied heavily on social constructivism, in form of qualitative data than numerical data. In conclusion:

- Sensitivity to participants' value system and knowledge and a need to validate or triangulate required co-constructing the process and its outcomes (Constructivism)
- The desire to make the evaluation a change agent meant that the process was developmental
- Considering that most of the opinions and acts are difficult to quantify meant that the data was qualitative
- Measurable attributes required quantitative processing (and statistical analyses)

I have to point out that different schools of thought might place what others see as philosophies under methodologies, i.e., the difference between research philosophy and methodology does not seem to be clear. Hence, I reference the placements.

a. Qualitative and quantitative evaluation

I indicate the activities, by which I identified the qualitative and qualitative data collection and analysis (Table 12). Using both in a single research is common (E.g., Madaus & Kellaghan, 1992: 127-134; Neuman, 1997; Guba & Lincoln, 1994; Heron, 1996: 160; Denscombe, 1998: 208; Savenye & Robinson, 2001: 1171-1172).

Table 12: Quantitative and quantitative procedures in this evaluation

What was done in this evaluation process	***Quantitative research***	***Qualitative research***
I started this evaluation with rationales, and sought for causal relationships. However, the events during the evaluation were in themselves data.	Hypothetical, and causal deductive	Meaning is derived from the research activity and patterns of behaviour
I did pre-planning, but I changed plans to suit the participants' convenience and circumstances. E.g., I changed dates or times when necessary.	Data collection is pre-planned and the plan is adhered with	Some methods of collection arise in the process of research
There were numerical data such as test scores, but these were accompanied by detailed descriptions of contexts and participants' interpretations.	Measure facts / numerical data	Construct social reality from observing events and from statements
I considered validity and context of utmost importance, and so worked with participants, listening to their interpretations of events	Reliability is important	Authenticity is important. Concern with meanings and understanding
I was sensitive to values, context, and culture of participants. E.g., I respected the culture of participants in relating with them and in terms of what they wanted to gain from the evaluation exercise.	Value free	Full of value systems
Unique contexts were considered independently, but common emerging themes were combined for statistical analysis	Independent of context, and can be replicated	Contextual
Constant comparative method led to themes, and discourse analysis sought participants' interpretations and the quality of interactions between participants and the programmes.	Statistical analysis	Thematic analysis
I was a facilitator of the evaluation, and a member of the evaluating team.	Researcher is detached	Researcher is part of the process
Selected schools represented disadvantaged schools in SA.	Sampling follows a design	Sampling is purposive – deliberate non random selection

For the purpose of this evaluation, I took qualitative research to comprise the nature of activities and realities from which information on how participants went about evaluating CAA and Zadarh was obtained. It was preferable for this evaluation to be qualitative to capture context-specific social transactions involved in evaluating programmes as well as participants' views (Guba & Lincoln, 1994; Madaus & Kellaghan, 1992: 133; Chua as cited in Myers, 2000: 2; Kuiper, 1997: 11; House as cited in Blasi, 1999; Elliot, 1991: 216; Madaus & Kellaghan, 1992: 133). Guba & Lincoln (1994) argue that qualitative approaches are sensitive to the process.

Of the many qualitative research categories (Greene, 1994: 532; Reeves, 2000a: 25; Reeves & Hedberg, 2003: 34), I deemed the developmental, post-modern, interpretative and constructivist approaches more appropriate. *These paradigms used in a qualitative framework are upcoming alternatives to analytic-empirical-positivist-quantitative paradigms* (Reeves& Hedberg, 2003: 31).

I have noted the counter arguments against qualitative evaluations, especially that they lack empirical proof. Hannafin & Hannafin (1991:304) mention *editorial bias against qualitative research.*

I propose that quantitative measurements too approximate human behaviour when scales or scores are allocated to an arbitrary 'feel' or an assumed act, without the input of the individual being observed. Hickey & Zuicker (2002: 541-542) discuss other concerns, including the bias in qualitative research due to value systems. Again, a quantitative measure of human activity is a value judgement; a researcher has to decide whether an act is valuable or worth for the research data.

Finally, the choice between qualitative and quantitative approach was influenced by the objectives for which CAA and Zadarh were designed, and what aspects of these programmes might interest end-users in. One of the objectives of this evaluation was to see how ECPs could change the teachers' classroom practices. Since C2005 is a constructivist curriculum, I was interested in the constructivist aspect of the two ECPs, which in my opinion would be better accessed through qualitative strategies, such as constructivism and interpretativism.

b. Post-modern evaluation

According to Nichols & Allen-Brown (2001: 231) and to Yeaman, Hlynka, Anderson, Damarin, & Muffoletto (2001: 254-256), postmodernism rejects 'modern' epistemological foundations such as positivism and also resists dominant, oppressive cultures, and shifts the power of seeking knowledge to the people. The understanding applied here followed the explanations given in Reeves & Hedberg, (2003: 271). This evaluation was post-modern because it was sensitive to the assumptions underlying the educational programs (i.e. whether they try to manipulate teachers) with the ultimate aims of finding out how and whether they are compatible with classroom practices, and whether the programmes help teachers to improve their classroom performances, and evaluation skills. In this way, the evaluation process was interpretative in relation to all participants, and would empower the teachers, considering that the majority of them have been disadvantaged in terms of using computer technology.

Furthermore, I set no strict rules and I interacted with participants at their convenience. I was part of the evaluation process and made sure that participants understood the process and how the programmes worked, by running the programmes with them. These measures helped me to understand the social dynamics in the schools, in relation to computer uses in schools and to evaluation.

c. Developmental evaluation

Two basic tenets of development research that were important in this evaluation were that the evaluation had to establish collaboration between the teachers, learners and me, and that the designers, the teachers, and I benefited from the evaluation process (Reeves & Hedberg, 2003: 275). These tenets imply that teachers had to learn, not only how to use the two ECPs, but also how to evaluate them. It was a process, which drew upon situated learning, interpretativism, and constructivism (see learning theories, Part II), whereby teachers learn in their schools and with their learners. Thus, the evaluation was valuable in the context under which teachers would use the two ECPs (Herrington, 2002). This development model was linked to an action methodology (explained later).

Another dimension of development was the feedback to instructional designers of CAA and Zadarh, who would hopefully consider the findings from the evaluation when they develop these two ECPs further. Furthermore, the evaluation was also developmental in terms of checking the applicability of the selected theories of learning adopted in the ECPs, theories of evaluation and of constructivism towards solving problems science teachers face in class as well as in evaluation. Both of these developmental facets hopefully contributed towards changes in the strategy teachers used in class and the way I, as a facilitator of the evaluation and participants understood the paradigms applied in the process.

The developmental approach validated the evaluation in terms of '*External value constraints'* since the evaluation *was valuable for informing and improving educational practice (*Eisenhart & Howe, 1992: 660), as well as in terms of '*Internal value constraints'* because the findings were authenticated by training participants in evaluation and about the programmes (Eisenhart & Howe, 1992: 661).

The process followed the advice as given in Reeves & Hedberg (2003: 274), except that this particular process was integrated or recursive

d. Constructivist evaluation

It is human nature to construct knowledge or to confirm held constructs from interpreting an experience. Hence, constructivism is recommended in science and in research on instructional design (Wilson, 1995a: 5; Cobern, 1996: 304).

I wish to reiterate that this research aimed at deriving a workable but useful evaluation scheme, while actually evaluating two computer programmes. I used constructivism in two ways: to map out the methodology of the evaluation process and; to obtain views on the value of the programmes. The evaluation scheme emerged from participants' advice on the best way to evaluate the computer programmes, as we went about judging the utility of the computer programmes. For example, during the evaluation process, I removed or rephrased questions, which participants persistently judged as being meaningless. The evaluations resulted from dialogue during which a participant and I would try to agree on interpretations or constructs (Guba & Lincoln, 1994: 111; Wilson, 1995a; Willis, 2000: 12; Winn, 1997), although I provided some guidance in forms of questionnaires and interviews (Willis, 2000: 12-13). That is, the evaluations and the construction of the evaluation scheme were negotiated and process-based in a way outlined by Cennamo, Abell, & Chung (as cited in Willis, 2000). Cennamo *et al.* advise that the process should be driven by social negotiations, be client-centred and should nurture reflexivity, permitting participants to ask questions rather than asking them to complete tasks. That is, no prescriptions were followed, except running the programmes.

Methodology

Introduction

A methodology describes the plan of action, process, or design behind the choice of methods (Crotty, 1998: 3). The aim of a methodology is

> *To describe and analyse methods, throwing light on their limitations and resources, clarifying their presuppositions and consequences, ... It is to venture generalizations from the success of particular techniques ...* (Kaplan, as quoted in Cohen & Manion, 1987: 42).

Cohen & Manion continue that the aim of methodology is to help us to understand the process of research.

Ethical considerations

Data collection techniques present ethical dilemmas especially because researchers invade the participants' privacy and so need to have a strict code of ethics (Merriam, 1998: 214). The first ethical consideration was securing permission from the DoE for the evaluation to take place in schools, and then permission from school principals to involve teachers and learners. The second consideration was that I did not reveal participants' identities and responses by using pseudonyms in this thesis. Thirdly, I shared the findings, and provided a copy of the evaluation scheme I developed to participants. Fourthly, I explained the study in a covering letter accompanying questionnaires and later demonstrated the programmes, and then asked for willing learner and teacher volunteers. Cates & Goodling (1997: 30) also used volunteers in a similar study. Volunteering implied that a participant was interested, and unfortunately eliminated those who would have negative attitudes towards the program at the beginning. Fifth, I avoided embarrassing and personal questions.

Action evaluation

If teachers are often unwilling to make use of technology due to anxieties and myth surrounding the computer (James, 1988: 190), this evaluation had to encourage them to remove these anxieties and myths. Stevenson (1995: 207) advises that such changes should evolve from within the teachers and the school. These considerations require action research, which *focuses on social concerns, values, and perspectives of participants and their context* (Holloway, 2001: 1109, 1124). Reeves& Hedberg (2003: 272) give an adequate coverage of action goals. One such goal was to enable teachers to use the two ECPs effectively, bearing in mind the underlying educational and science theoretical frameworks, to increase their understanding of evaluation, and in the process to change their classroom strategies.

I implemented action evaluation by running the ECPs with teachers to remove myths and increase their confidence, and through further appointments (for example, when I went to collect completed questionnaire), by helping teachers to run and use the ECPs for their learners. That is, I left the CDs of the programmes with willing teachers, after helping them to install the programmes on their computers so that they, and their learners could familiarise themselves with those programmes, and encouraged them to telephone me whenever they needed clarification or help. I also designed a manual for Zadarh basing it upon the questions that were frequent from learners and teachers and gave it to them. Question Mark had a website from which I downloaded the user manuals. Finally, I attended some of the sessions when learners were using the programmes, and assisted teachers as learners were using the programmes. I repeated the attendance when possible, on account of Noffke's (1995: 4) advice that understanding and actions emerge in a constant cycle, with educators needing constant support.

Interpretative evaluation

Qualitative evaluation can be interpretative (Myers, 2000; Neuman, 1997: 329; Schwandt as cited in Heron, 1996: 160; Greene, 1994: 532), because interpretativism rejects the view that anyone's creation is the way to think, and considers knowing as a local event (Wilson, 1995a: 5; Reeves & Hedberg, 2003: 28). I used the interpretative approach to attain the subjective and contextual qualitative interpretations that individual participants gained from experiencing the programmes (Cohen & Manion, 1987: 36; Greene, 1994: 536; Heron, 1996: 160; Reeves & Hedgerg, 2003). That is, I obtained data by *Versteben:* That is, by subjective, participative understanding without verification to external criteria (Madaus & Kellaghan, 1992: 133), and I analysed the individual interpretations with regard to the context and participant profile. Seidman (as cited in Hitchcock & Hughes, 1995: 295-304) supports this approach in the statement that:

> *"Instead of appealing to absolutist justifications, instead of constructing theoretical logics and epistemic casuistries to justify a conceptual strategy, … we be satisfied with local, pragmatic rationales for our conceptual (interpretive) approaches".*

These local rationales include increased possibilities of understanding between participants.

Although there is richness in separating data from different contexts (Greene, 1994: 536), problems with interpretativism emanate from accepting subjective contextual knowledge as reality (as opposed to scientific objectivity) (Guba & Lincoln, 1989: 44). Interpretativism precludes common standards for evaluating information, and holds that no information is universal (LeCompte *et al.,* 1993: 324-325). I enhanced generalisation, objectivity, and reliability by carefully describing the profiles of each participant (LeCompte *et al.*, 1993: 333), and the school contexts under which participants operated. For example, descriptions included qualifications and experiences. I sought objectivity by a constant comparative analysis of statements made by participants or of the transactions between participants and me (Greene, 1994: 536), from the points of view of the participants (Myers, 2000: 2), which I considered to be social constructions (evaluations).

Participation

Teachers and learners evaluated the programmes (Table 13) to ally fears enumerated by the National Research Foundation (NRF) (2004) and Elliot (1991: 217), and following guidance from Reeves' (1994) and James's (1988: 193-195) models. A major concern, also considered by Miller & Olson (1994) and Hitchcock & Hughes (1995: 32) is that the integration of ECPs in schools should involve teachers. Teachers are the implementers of programmes, and their experience in the evaluation could help them to evaluate other programmes and to incorporate the ECPs in the curriculum successfully. The number of computers and of volunteers determined the number of participants in each school.

Similar to Senior (1990: 61), I considered learners' opinions about the programmes imperative. Learners' views tend to be more accurate than parent or teacher responses (Shakeshaft, 1999: 4). Although Barbera (2004: 15) denounces that consideration, in this case, it was important to evaluate, *inter alia*, learner motivation and participation in own learning, as well as the processes of learning programmes promote.

Table 13: Roles of participants

	PROCESS	TEAM	PRODUCT
ANALYSIS	o Needs analysis to find out whether the programmes could be useful o Planning the research and evaluation strategies on the basis of contexts, with participation of teachers Finding out the level of understanding that teachers had about: Evaluation and the use of ECPs; Subject content; and Learning theories	Researcher Teachers	o Needs analysis o Identifying participants o Teachers' profile o Validation of programmes in terms of content and approach o Evaluation Plan
EVALUATION	o Testing or piloting the computer programmes o Focus-group evaluation (learners) o Analysing results	Researcher Teachers Group of learners	o Evaluation report and recommendations o An evaluation model for interactive programmes

Case evaluations

Qualitative case evaluations are common in information systems and include interpretativism (Myers, 2000; Greene, 1994: 532), because cases highlight the differences among participants and contexts in terms of the way they respond to a programme (Greene, 1994). Together, cases provide enough information for establishing generalisations (Cohen & Manion, 1987: 120; Wolcott, 1992: 28; Campbell in Stake, 1994: 238).

The uniqueness of these formerly disadvantaged schools in SA required the evaluations to be case studies. These are for example socially and economically different from schools in SA, which manage to obtain computers by themselves. However, I had to constantly monitor biases I could make in my judgements due to the uniqueness of the cases. Furthermore, the study was limited and unique by subject area (science), by context (ID in South African schools) (Stake, 1994: 236; Wolcott, 1992: 28; Cohen & Manion, 1987: 120; Hitchcock & Hughes, 1995: 319), by the sample of schools that had computers (Huysamen, 1994: 168), and by the two computer programmes that were evaluated.

Bassey (as cited in Bell, 1992: 9) stated that an important criterion for judging the merit of a case study is the extent to which the details are sufficient and appropriate for a teacher working in a similar situation to relate his decision making to that described in the case study. The main advantage in approaching the evaluations as case studies was that I obtained more detailed in-depth understanding of the impact of the two ECPs upon participants in the context of science education in disadvantaged communities (Greene, 1994: 532; Bell, 1992: 8; Neuman, 1997: 331; Elliot, 1991:216; Imenda & Muyangwa, 1996: 30). These cases provided opportunities to refine (Stake, 1994: 237) earlier evaluation schemes, while empowering targeted participants (Hitchcock & Hughes, 1995: 322) with skills of using ECPs, with constructivist approaches to learning and with deeper conceptual understanding.

That these were case studies also required specific planning whereby each school set an agenda for the evaluation suitable for its timetables and availability of computers. The planning went as follows:
Munro (1975: 221) points out that valid curriculum evaluation can only be made in a typical classroom atmosphere. This point implies a number of considerations, two of which are that the evaluation considers the nature of the subject, and the strategies recommended to teach that subject. Hence, I started the evaluation of the two ECPs by a review of the theoretical foundations of science and learning, as well as preliminary study of the problems in some of the schools that participated. Furthermore, I carried out the evaluations in school computer laboratories.

I initially made appointments by telephone, and then visited interested schools, during which I explored the availability of a computer laboratory, the interest and willingness of teachers and learners to participate. These initial visits involved validating the computer programmes against participants' needs (Reeves & Hedberg, 2003: 119; Shakeshaft, 1999: 3), and piloting the research tools on a few teachers and learners. I then made further appointments for the study during that visit.

In consideration of the different disciplines in ID (curricula issue, science, and computer technology), the different interests of participants (Percival & Ellington, 1984: 118), I had to involve experts from each of those disciplines. For example, subject advisors represented government positions on curriculum issues and problems, I am a qualified scientist and looked at how programmes represented science, and an ID expert gave opinions on the design of the programmes. I looked at these opinions against the philosophies of science and against the learning theories, whose literature I provide in the next two chapters.

Discourse analysis – did participants mean what they said or wrote?

I adopted Cohen & Manion's (1987: 253) definition of discourse analysis as an examination of 'accounts as they occur in context, bearing in mind the definitions of discourse such as: *'Communication in speech or writing' .. usually serious, subject'* (Cambridge International Dictionary of English, March, 1999); ... *a form of power that circulates in the social field and can attach to strategies of domination as well as those of resistance' (*Diamond & Quinby as cited in Pinkus, 1996*); Any structure of knowledge which determines the way in which the world is experienced and seen'* (Collin's Dictionary of Sociology 1991: 196*)*. Foucault offered guidance on the constitution of the discourse, which I took into account (E.g., in Mphahlele, 1996). Foucault pointed out that discourse (in form of text, disposition, experiences, statements, speech, actions, objects in the environment, gestures, glances, attitudes, thoughts, understanding, ideology, values, and emotions, etc) is a way of constituting knowledge, and that, together with the social practices, explicates power relations, and influences the way people are governed (Pinkus, 1996). That is, power can shape, govern, and structure the discourse to the extent that an evaluator can misinterpret data, if the discourse in not taken into consideration.

In the cultures in which this evaluation happened, there is often a hierarchy of power and respect, which is an important dimension in the administrative structure in schools. For example, an elderly teacher or royal blood is bound to be respected by even those professionally above. Furthermore, teachers might decide to agree with everything I say because I am bringing to them technology from which they think they will benefit. While much of the power relations were beyond the influence of the evaluation process, I had to be conscious of relations within schools, and between participants and myself. Discourse analysis provided the contextual and participant value landscape upon which I analysed and validated the emergent constructions.

In the context of using English, which was a Second Language to participants, and where computers were new innovations, I believe that it is compulsory to analyse the discourse because participants might have used some terminologies inappropriately. For example, 'easy' could mean the same emergent theme of 'fair' depending upon a participant's disposition at the time of using the programme, and could be related to the participants' anticipation to gain from the evaluation. I tried to remove barriers, for example, by use of simple English that was in reach of participants, and sometimes re-phrasing questions or obtaining a vernacular translation, and had to work together with the participants, who actually became co-evaluators.

Methods used in this evaluation

Introduction

Cohen & Manion (1987: 42) define a method as *that range of approaches used in research to gather data, which are to be used as a basis for inference and interpretation, for explanation and prediction*. Methods are *simply procedural aids for testing the methodological conceptualisations of the investigator* (Altrichter, as cited in Elliot, 1991: 218). Thus, methods move from *philosophical assumptions to research design and data collection* (Myers, 2000: 5), where a research design is a the plan, which includes how research participants are obtained and what is done to them *with a view to reaching conclusions about the research problem* (Huysamen, 1994: 20). Reeves (2000a: 23; 2000b: 7-8) advises that the objectives and tasks of research determine the methods and methodology. In these cases, the objectives was to see that participants gained from the evaluation process, and so methods had to help them gain and at the same time reveal what and how they had gained.

I used assessment, questionnaires, observations of users, and interviews in an interpretative approach (Bell, 1992: 8; Greene, 1994: 532) that gave participants a platform to be critical of the ECPs and the evaluation process. The participants and I also considered cost and convenience (Heinecke, *et al.*, 1999: 3). The following section gives the details of methods and the methodology applied in this evaluation.

Questionnaires

Among the several definitions of a questionnaire, I found that of Galfo (as cited in Sanders, 1995: 712) most applicable: *data-gathering instruments used to obtain factual data, opinions, and attitudes in such a way that the respondents and the data-gatherer need not come into contact with each other*. Questionnaires comprised a list of questions administered on paper. The design of questionnaires followed advices from various sources such as Cohen & Manion (1987). For example, I made questionnaires attractive (by spacing and using clear fonts), gave clear instructions, and sequenced questions logically by attending to a particular issue at a time.

For the main advantages of obtaining data from large numbers of respondents, I applied questionnaires to subject advisors, teachers, and learners. I reduced the disadvantage of misunderstanding questions (common among those who use English as a Second Language), and of possibilities of untruthfulness, by going through the questionnaires with the participants, and asking them whether they understood every question. I also encouraged them to keep their unanimity and of responding when it was convenient to them although I preferred to be present when they were filling up these questionnaires. My presence enabled interviewing respondents immediately on matters of interest among their answers.

Interviews

I considered an interview to be a conversation or dialogue with the purpose of exchanging information, thoughts, feelings, values, and understanding between two people (Cohen & Manion, 1987: 291-293; Merriam, 1998: 71; Wolcott, 1992: 28; Hitchcock & Hughes, 1995: 154; Soriano, 1995:19-21; Clancey as cited in Wilson, 1995b; Witkin & Altschuld, 1995: 145; Fontana & Frey, 1994: 361-364; Denscombe, 1998: 109-112).

Some interview questions explored responses participants had given in questionnaires. These were only necessary where responses in questionnaires were ambiguous or interesting in some way. Other interviews happened at the time participants were using the programmes. We used think-aloud protocols at the time learners or teachers were using the ECPs. Then participants were interviewed to give a summative view of their experiences with the ECPs. These summative interviews highlighted events that participants felt important. Interviews were generally about the participants' experiences with the computer programmes and also about the evaluation process.

Follow-up questions, either after responses in the questionnaires or in the interviews were phrased to help a participant to reflect further on his/her answer – that is some questions were reflexive (Willis, 2000: 9; MacDonald & Farres, 2003: 5; Willis, 2000:

9; Guba & Lincoln, 1989: 44). I used telephone interviews when it was difficult to meet participants.

Fortunately, participants cooperated and were motivated, thus meeting the concerns Kitwood (as cited in Cohen & Manion, 1987: 294) raises. Participants were motivated by the opportunity to learn something related to computers such that, I think that their conversations were genuine since they continued to telephone me as they were using the programmes.

Amir & Tamir (1994: 94) suggest that *descriptions of misconceptions are very important starting points,* and Gredler (2001: 538) advises the evaluator to *request learners to verbalise their thoughts as they work.* I prompted learners to make verbal predictions about a certain activity, to describe what they observe when performing the activity, and to compare their predictions to their observations. I interjected with open-ended and diagnostic interview questions to expand on interesting interviewee statements or misconceptions, and I corrected misconceptions during the discussions. These approaches are diagnostic and reflective, and helped to identify solutions and problems programmes provided, and how they assessed the evaluation (Goldstein as cited in Little & Wolf, 1996: xiii; Linn, 2002: 40).

Therefore, the interviews were semi-structured (as defined in Cohen & Manion, 1987: 293, Hitchcock & Hughes, 1995: 159; Denscombe, 1998: 113; Merriam, 1998: 74; 1987: 293; Bell, 1992: 91; and Soriano, 1995: 20) in the sense that there were pre-planned questions, but these were not presented in a strict order. I revised a question or translated it into vernacular when I saw that the respondent did not understand the question.

Furthermore, I improved upon questionnaires and interviews as the research proceeded using the questions the participants asked me. This procedure helped me to fine-tune the clarity of questions and to improve upon the evaluation scheme.

Focus group interviews

Participants are more responsive and assist each other in constructing a sense of their experiences when they are grouped (Fontana & Frey, 1994: 364; Cohen & Manion, 1987: 294; Millward, 1998: 276; Witkin & Altschuld, 1995:171; Queeney, 1995:124-129). Furthermore, Soriano (1995:23) explains that the *participants selected, determine the dissemination of the group's responses to broader populations and, that they work best when views of homogeneous sub-populations are sought.*

I used focus group interviews of not more than five learners to obtain in-depth cumulative data quickly in schools where for example teachers wanted the sessions to end. This involved discussing one question at a time with all participants in a group. I interviewed females and males separately, to tease out any gender tendencies, since females and males tend to play different games or to desire different activities.

Participant observation

Observation is common in evaluating multimedia interactive learning resources (Huysamen, 1994: 169; Various research projects at the BECTA, 2001). Such measures provide a means of capturing the interactivity engendered by these media.

I photographed or video-recorded the participants where possible, and analysed video records against the interview and questionnaire data. I used Observation Schedules (Science Teaching Observation Schedule [STOS], and the Play Observation Schedule [POS]) as additional sources of information. These schedules were modified from the STOS used in Muwanga-Zake (1998), which was found to be informative about classroom proceedings. A reflexive dialogue at the time participants were using the programmes sometimes accompanied the observation, especially when I sought to gauge enjoyment.

Assessment

Another source of data were diagnostic assessment items, some in the traditional test format, while others were applied through reflexive dialogue. Results from learner assessment can be used to evaluate instructional programmes (Percival & Ellington, 1984: 110-120; DoE, 1998; Madaus & Kellaghan, 1992), for example, by use of tests. Besides, *assessment lies at the core of learning* (Lloyd-Jones *et al.*, 1986: 1). Other authors show similar conclusions (E.g., Harlen, 1991: 325; Harlen, 2000; Linn as cited in Madaus & Kellaghan, 1992: 126; Lawton & Gordon, 1996: 15; Little & Wolf, 1996: ix; Munro, 1975: 222 - 223). The reflexive dialogue helps to give ownership to participants if it is combined with an interpretive paradigm, countering the observations in Madaus & Kellaghan (1992), that assessment has been used in evaluations by the powerful to define how things are.

I viewed assessment items in light of learning theories. This is because, according to several theorists (E.g., Madaus & Kellaghan, 1992; Gipps, 1996: 251 - 252; Ertmer & Newby, 1993), assessment should be positioned in a learning theory. For example, within the constructivist paradigm I used, learners were allowed to ask questions during the time they were doing the tests on CAA or while they were playing Zadarh such that assessment was *integral* with *learning* (Natal College of Education, 1997: 109; Black, 1986: 7-8; Lloyd-Jones *et al.,* 1986: 2; Harlem, 1993: 159-168; King & van den Berg, 1992: 17; Pausch & Popp, 1997: 1; Harlen, 2000). I assessed learning obtained from Zadhar by sitting with the learners as they completed tasks in Zadhar, and then asking them about what they valued and learnt from Zadarh. I also encouraged teachers to accommodate learner's questions. Teachers too asked me questions for guidance and on concepts they found problematic.

The assessment was qualitative because it sought the participants' opinions, interpretations, and constructs using open-ended questions. This kind of assessment was developmental since it judged where learners were in their understanding (Pausch & Popp, 1997: 1-2). I, however, note Hein & Lee (2000: 5) advise that using assessment can be dangerous if one is evaluating teachers, since assessment can leave out other important factors that affect learning. I considered some of the other factors by using other methods besides assessment.

Data Management and Analysis

Introduction

Several workers advice researchers to give an account of themselves in the research and of the analysis techniques (Gay & Airasian, 2000: 204-211; Huberman & Miles, 1994: 428; Denzin, 1994: 502; Hitchcock & Hughes, 1995: 295-304; Altheide & Johnson, 1994: 493-494; Maykut & Morehouse, 1994:145-147). I include my deeds as part of data besides the data I collected on and from teachers and learners. Furthermore, QM automatically analysed some of the assessment data. For the rest, the philosophies, methodology, and assumptions informed analysis of data.

Although Denzin (1994: 502) warns researchers from masking themselves behind paradigms, I analysed data from an idiosyncratic constructivist and interpretive approach, but separated my opinions and biases from participants' data. Discourse analysis, by which I clarified the meaning the participants attached to words in light of English being a second language to them, and myself, as well as giving this document for editing by experts in English, helped to reduce bias further. Furthermore, I reduced bias by coding participants (i.e., hiding their real identities). The code L1 meant learner 1, and T1 meant teacher 1. In this case, I took bias to be a systematic or persistent tendency to make errors, by, for example, overstating or the true value of a programme (Cohen & Manion, 1987: 302; Bell, 1992: 91; Soriano, 1995: 20-21).

Methods adopted in the analysis of interview and questionnaire responses

I focussed upon participants' comments in relation to benefits or problems (Percival & Ellington, 1984: 110-120). I also analysed the context and emerging constructions participants made using constant comparative methods, and I supported these by statistical analyses (E.g., Pearson's' correlations) and review of video records. Why did I choose these methods?

a. Constant comparative method

I used a constant comparative method to analyse some of the participants' especially for similar questions, to determine recurring themes and patterns in participants' responses as suggested by Maykut & Morehouse (1994:126-144), Hitchcock & Hughes (1995: 295), and Denscombe (1998: 210) as follows.

- Statement on event / ideas and suggestions
- General theme or unit of meaning. E.g., exciting
- Underline words or statements related to exciting and code them with a number: E.g., makes me feel good', 'not boring', I can go on and on', 'makes me happy', all belonged to the general unit, 'exciting'.
- Compare unit: E.g., compare 'exciting' against closely similar themes such as 'motivating'. These were combined if their meaning overlap to form a new unit. Otherwise, each unit was refined where the two were definitely different. This step required setting rules of inclusion for each unit.
- The rules of inclusion described the unit.

I continuously refined these units, each with an audit trail. I then derived generalisations from each unit.

b. Statistical analyses

I used SPSS to analyse some of the data by using correlations to check on the significance level of some relationships. I used the tool to produce graphs where necessary.

c. Review of video records

I used video reviews to confirm my observations, the discourse, and themes, which had emerged.

The aspects of the programmes that were evaluated

I followed advice from several sources such as Castellan (1993: 234 – 235) and Percival & Ellington (1984: 110-120):

- *Technical accuracy – whether programmes execute easily, correctly and accurately on school computers*
- *Pedagogical soundness – whether instructional goals are articulated clearly, it is clear where technology ends and subject begins, technology is appropriate to the concepts learned, learners can use the programmes, and whether the technology permits self-assessment, and to what extent the programme covers or fits into the curriculum.*
- *Substantive fidelity – concerns accuracy of content and whether the content is worth learning.*

The evaluation was limited to Kirkpatrick's level 1 (reaction - the acceptance and use of the programs) and level 2 (the analysis of learning from the programmes) (Boverie, Mulcahy, & Zondlo, 1994: 82). In the context of a developing community, I began the evaluations by finding out the problems in science education, validating the programmes against needs (Reeves & Hedberg, 2003: 119; Shakeshaft, 1999: 3), and checking on the capacity of both human and technical resources at schools.

Problem-solving

After reading through literature, for example, in the above, one recurring concept is problem-solving. Problem-solving appears as a scientific process, and as an essential ingredient in cognitivism and constructivism. It is therefore not surprising that problem-solving features prominently in CAA and Zadarh.

Jonassen *et al* (2003: 20-29) debate problem-solving in detail. In the context of CAA and Zadarh, one concern that requires attention is how well defined problems were. Jonassen *et al.* give advice towards this concern: They state that well-structured problems present all elements of the problem, engage a limited number of rules and principles, have preferred, prescribed solution processes, and possess correct, convergent answers. Such problems seem behaviourist (as has already been argued in behavioural assessment). On the other hand, ill-structured problems may have many alternative solutions – these are likely to be subjective and constructivist. On addition though, ill-defined problems can be vaguely defined, with no clear goals and constraints, and have multiple criteria for assessing their answers.

Although I have deliberated upon these issues at length in the section on assessment, the evaluation of CAA and Zadarh ought to carefully look at how the problems are given and assessed.

Conclusion

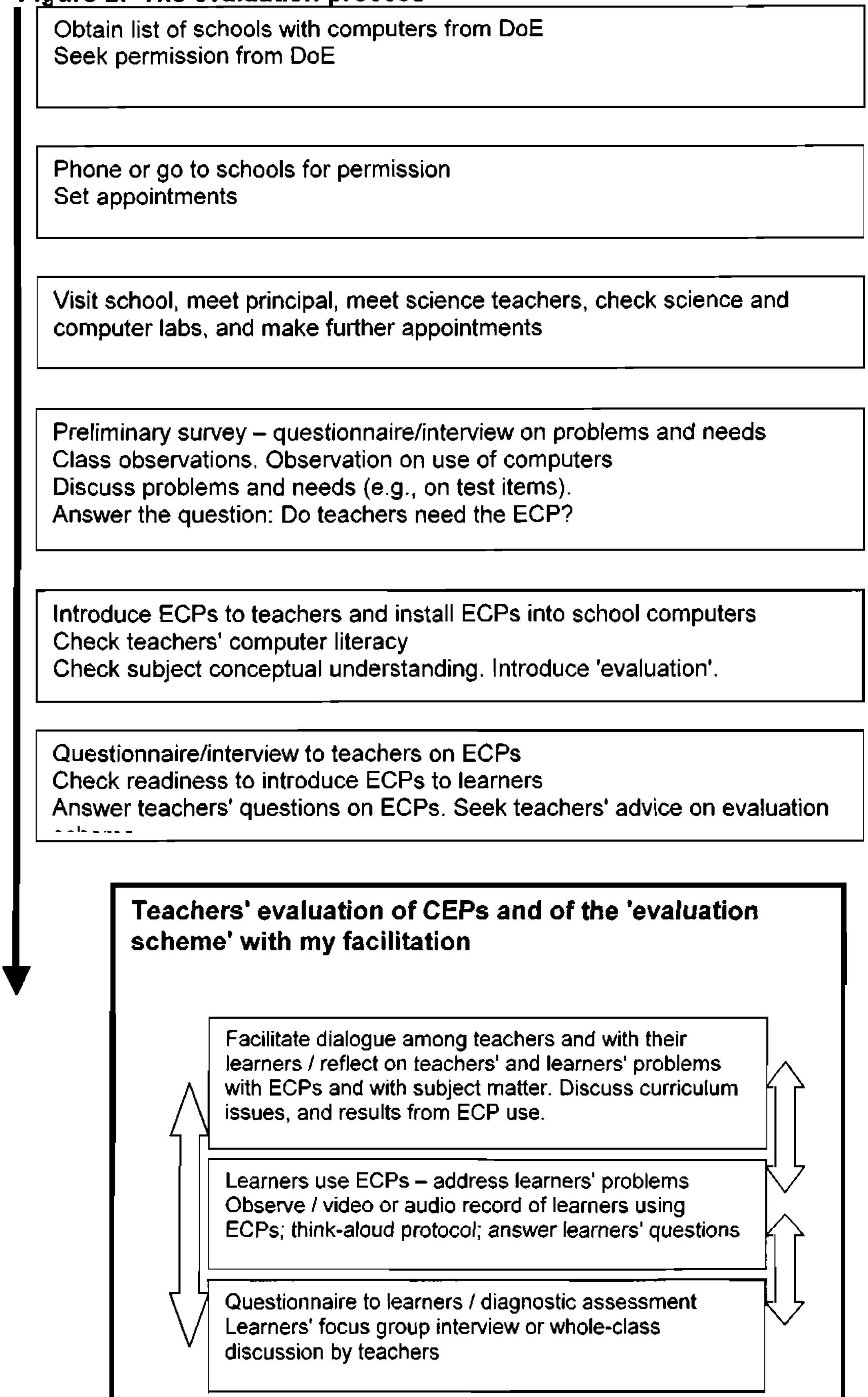
I designed this evaluation to be one of the teachers' activities as they use computer-aided educational programmes. The activities were developmental, and changed the teachers' practices positively by helping teachers to identify problems in their classrooms, think about, and seek solutions, and to lead them towards better classroom practices. I have summarised the philosophies and methodology I used in Figures 1 and 2 on the next two pages.

Therefore, the next two chapters give an outline about the theoretical considerations in science and education, followed by how these theories might relate to the problems found in science classrooms.

Figure 1: Conceptual framework of the evaluation process

Philosophy	Methodology	Methods	Actors / Audience	Data	Data analysis
Qualitative *Teacher experiences not quantifiable* *Some experiences were unique in each school or to individual teachers and learners*	**Case study** *Differences in contexts noted* *Specific to science, Grade 10-12, disadvantaged schools in SA* *Specific to CAA and Zadarh*	Questionnaires Interviews (focus groups) Class observations Assessment	Teachers Learners Myself	Interview scripts Class notes Video or audio records Questionnaire responses	Constant comparative Discourse analysis
Developmental *Teachers gain from the evaluation process: evaluation skills; conceptual understanding; computer skills*	**Action** *Enable teachers to use ECPs and to evaluate by participation in the whole process*				
Post-modern *No rules – no formalities* *Sensitivity to compatibility of ECPs with curriculum and teachers' capacity*	**Class visits** *Observe teachers' practices* *Observe the way learners interact with ECPs and their teacher* *Opportunity for interview while using ECPs*				
Constructivist *Teachers and learners contribute towards the evaluation process* *Participants' questions inform the evaluation scheme*	**Ethical considerations** *Permission from school authorities* *Evaluation happened at teacher's convenience* *Unanimity in records*				
	Interpretive *Teacher's subjective and contextual interpretations recorded*				

Figure 2: The evaluation process



Obtain list of schools with computers from DoE
Seek permission from DoE

Phone or go to schools for permission
Set appointments

Visit school, meet principal, meet science teachers, check science and computer labs, and make further appointments

Preliminary survey – questionnaire/interview on problems and needs
Class observations. Observation on use of computers
Discuss problems and needs (e.g., on test items).
Answer the question: Do teachers need the ECP?

Introduce ECPs to teachers and install ECPs into school computers
Check teachers' computer literacy
Check subject conceptual understanding. Introduce 'evaluation'.

Questionnaire/interview to teachers on ECPs
Check readiness to introduce ECPs to learners
Answer teachers' questions on ECPs. Seek teachers' advice on evaluation

Teachers' evaluation of CEPs and of the 'evaluation scheme' with my facilitation

Facilitate dialogue among teachers and with their learners / reflect on teachers' and learners' problems with ECPs and with subject matter. Discuss curriculum issues, and results from ECP use.

Learners use ECPs – address learners' problems
Observe / video or audio record of learners using ECPs; think-aloud protocol; answer learners' questions

Questionnaire to learners / diagnostic assessment
Learners' focus group interview or whole-class discussion by teachers

PART V

TEACHERS' AND LEARNERS' OPINIONS ABOUT CAA AND ZADARH

Given the contexts (problems in science classrooms and utility of computers) outlined in the previous chapters, what do the potential end-users say about two computer educational programmes?

Introduction

In Part I, preliminary studies had indicated that science teachers faced problems in setting diagnostic MCQs. They also had problems in analysing and interpreting results, especially for in a diagnostic sense because of, *inter alia,* the large numbers of learners, and many class periods per teacher. These affect learning and contribute to higher failure rates by disabling frequent diagnostic assessment. Each individual learner should preferably be diagnosed, and frequently to improve performance (Butler, 2003), but diagnostic procedures require time for interviewing and correcting mistakes of each learner, as well as arranging some self-directed activities for the rest of the learners in the class (Bright, 1987).

This study

This section is a report on the evaluation of CAA at two schools in East London, Republic of South Africa during 1998 and 2001, involving two Grade 10 science teachers and 27 learners.

School 1 was a private school whose curriculum focussed on science, mathematics and technology. It was situated close to a main street through the East London central businesses. The teacher who participated was the head teacher for science, and she was also the science teacher for Grade 10. She was a 'Coloured' studying towards a Bachelor of Education (Honours). It appeared as though she had a good relationship with her learners. All the learners were Black.

School 2 was a government formerly whites-only school. Its curriculum was technical: E.g., it offered courses on electrical engineering. It was situated in an up-market suburb of East London, and the learners were of mixed colour, although it appeared to have more Whites. The teacher was a white male and did not reveal his qualifications.

The research questions

The main question of concern was: **How well can computer-aided assessment (CAA) be used to provide teachers with diagnostic information on learners'**

current understanding of the science concepts at grade 10 level? I unpacked this question into the following subsidiary questions:

i. **How technically sound is the test?**

Technical validity in this case refers to the appropriateness of the technology used in presenting the test and to the design of the Computer-Aided Test (CAT). It includes factors that could affect a learner' performance such as screen capacity, graphics, method of responding, and time limits (Sandals, 1992: 75-76). For example, the CAT was technically valid if it afforded the learners freedom of starting with any question, and if learners considered the colours friendly. The Technical Validity of the CAT had to be evaluated before the CAT could be evaluated for its diagnostic potential, since technical factors can affect the way learners engage with the CAT. The ID expert examined technical validity, and the questions I put to him appear in Appendix II.

ii. What is the quality of data that the CAT provides?

The quality of data that the CAT provides was examined using three questions (Table 14).

Table 14: Questions, methods used, and respondents used to answer the questions for evaluating the diagnostic value of the CAT

Questions	**Data collection methods**	**Respondent**
i. What are the learners' results on the test?	Test results	Learners
ii. What information does the CAT provide that is useful for diagnosing learners' knowledge? iii. How well do the results indicate the problems that learners have with the topic tested by the CAT?	Interview (semi-structured)	Teacher

Methodology

A science teacher and I had already set a test comprising multiple-choice items, which we typed into Question Mark Designer (Designer) to produce a Computer-Aided Test (CAT) (Test 2). (The teachers had set Test 1 before without my assistance). See Appendix I for Test 1 and Test 2.

Test 2 comprised 15 questions on Grade 10 Higher Grade, General Science, which had been validated by teachers, subject advisors, and learners in a preliminary investigation. It should be noted that QM has a licensing mechanism called a 'dongo', which is connected on the printer port of the computer, before one can type the test into QM Designer.

I then took the CAT to an ID expert for validation. After the teachers and I discussed and effected the correction that the ID expert wanted, we tried out the CAT and made sure that it worked, and that teachers knew how to run it. We set QM Designer with the following options:

- Question feedback: No immediate feedback.
- Question sequence: Learners were free to start with any question, or skip questions and could browse through the test back and forth, including the freedom to change answers.
- Captions: The screen showed the user's name, the name of the test, number of questions attempted so far, and time left.
- Other settings: A time limit of one hour, a user could escape from the test at any time, and answers were automatically saved upon the hard drive.

We then allowed learners to try it out in a school computer laboratory in which computers were connected to a Local Area Network (LAN), and asked them what they liked or disliked about the CAT during the time they were doing the test and after they had finished doing the test.

The learners wanted the CAT to show them their marks immediately, and the correct answers. Then we set QM Designer in such a way the CAT provided that information.

After a second round of learners doing Test 2, we looked at data that QM Reporter provided, and decided to focus upon Ohm's Law, which appeared to be among the problematic topics to learners (see QM reports below). We set a third test (Test 3) that was diagnostic on Ohm's Law and then learners did this test. This CAT retained the QM Designer settings.

The level of conceptual understanding required in Test 3 appeared high to learners (and to the teachers) that we had to discuss it with the learners as they did the test. We also set QM Designer with the option that allowed learners to attempt the test any number of times. Then we left learners to do the CAT by themselves. The teachers and I set Test 4 to be more diagnostic and left this on the school LAN for learners to do. All tests were left on the school LANs.

Methods

We obtained reports from QM reporter, and then I discussed the reports with the two teachers. I then interviewed the teachers about the data, their experiences with the CAT and on how they could use the QM reports.

Findings and analysis of findings

The time I spent on the study

It is important to point out that the whole process, from validating the CAT and getting learners to work through it took a whole school year; it took over 2 years if I include the time spent on securing permission from the DoE and then the school authorities, as well as the preliminary investigations and validations by the subject adviser. This is along period, in my opinion, but delays were caused by:

- The difficulty in getting permission from the principals to use the LAN in their schools

  I had to make appointments to meet the principals because they were not replying to my letters asking for permission.

  One of the schools used its computers to teach computer skills to nurses from a neighbouring hospital. The money from this venture apparently helped towards maintaining the school LAN. Therefore, permission had to be negotiated amid tight time schedules of the school LAN.
- Difficulty in finding an ID expert

  ID experts are very scarce around East London, and the one who agreed to help me out was often very busy.
- Lack of time on part of the teachers to set tests, look at the CAT, analyse data, and be interviewed.

  These teachers rarely had a free period to indulge into research and innovations. Secondly, they appeared to have problems with some concepts.
- Difficulty in getting learners to do the CAT.

  At first the schools looked at this exercise as an extra activity and not part of the school curriculum so much that learners did the first CAT during break or in the afternoons.

Technical validity of the CAT (see Appendix II)

The teachers found no problems with the technical aspects of the CAT (Answer to questions 5, 7 and 13, Appendix III). Besides their judgement, an Instructional Designer expert determined the Technical Validity of the CAT, and his judgement appears in answers to an interview (in Appendix II).

a. Loading the CAT

The ID expert did not find problems in loading the CAT (Answer 1a), and stated that all questions were easily accessible (Answer 1b). The expert also found that the CAT run efficiently (Answer 2 & 3). However, there were no dialogue boxes to help in case of need (Answer 6), and the programme filled the whole screen thus offering no room for accessing other computer activities (Answer 7).

b. Design considerations

The whole programme should be accessible and a user should be able to get in and out of any part of a programme easily. Therefore, many of the technical questions relate to the design of a programme (Questions 8 to 22). However, I describe those that might have impacted upon the performance of the learners.

The ID expert reported that it was not possible to get lost in the CAT as there were buttons that showed options (Answer 8 & 9). The CAT also indicated the status and questions a learner would have done, as well as time spent (Answer 10). The use of the mouse to answer questions was appreciated (Answer 16), and language in the CAT was also acceptable (Answer 19).

However, most of the concerns the ID expert raised appear in Appendix II as additional comments. These include the positioning of the text to the extreme top of the screen and lack of graphical representations, which are important in science. These were problems caused by limitations in QM Designer. The version we used allowed a limited space, which meant that long questions fill it up completely. This version does not easily import graphics or allow drawings.

The other problems the ID expert noted were decisions of choice. For example, the ID expert complained of an 'OK' button whose position was erratic, the use of different fonts, and the use of many colours. These were deliberately designed to make learners more interested and to make them aware of the different questions. The ID expert complained about errors in the Test 2, which no other participant identified.

c. Printing

The ID expert noted a problem with printing (Answer 23). The test could not be printed apparently because of the 'dongo'. So we could not obtain a copy of the CAT as it looked on the computer screen. The only printings possible were the reports, and that was after closing the CAT.

Data produced by QM reporter

QM reporter produces a number of reports, but I focus on those that are relevant for diagnosing learners' problems. These include response analysis, individual learner, summary, and list reports. I give examples below.

a. An example of a response analysis by QM (Test 2)

I have used as an example, Question 6 to present the kind of information that QM Reporter provides about how the class has performed at a question.

Question 6. Multiple Response - " Select THREE factors, which influence electrical resistance from the list below;" The following is the report on the way learners performed on question 6

Number of times question answered: 14 Average score: 3.43;

Maximum: 6 Minimum: 2 Standard deviation: 1.22

% of learners	**Choice**
21%	"3 items selected: Diameter, Type of material, The strength of electrical current"
7%	3 items selected: Temperature, Diameter, Type of material"
29%	"3 items selected: Temperature, Mass, The strength of electrical current"
29%	"3 items selected: Temperature, Type of material, The strength of electrical current"
7%	"3 items selected: Type of material, Mass, The strength of electrical current"
7%	"3 items selected: Temperature, Diameter, The strength of electrical current"

Note that only 7% got this question correct, and that overall the rest (i.e., 93%) believed that 'the strength of current' influenced resistance.

The Response Analysis report also gives a measure of 'facility' and of 'discrimination'. For question 6 above, Facility was 0.57, and Discrimination was 0.70. (Figure 3).

Figure 3: The meaning of facility and discrimination (as described by QM)

> **Facility**
>
> This is the level of difficulty of the question, from 0.0 to 1.0. It is calculated as the average score for the question divided by the maximum achievable score (assuming this is greater than zero). A facility of 0.0 means that the question is very hard (no-one got it right); a facility of 1.0 means that the question is very easy (no-one got it wrong). If you are measuring performance, questions are most useful if they are neither too easy or too hard. A very easy question is answered correctly by everyone, and so doesn't help you measure performance. Similarly a very hard question is answered correctly by very few people, and so adds little to your measurement. An ideal facility is 0.5, which roughly means the question is answered correctly by half the users. You may want to look carefully at questions that have facility more than 0.75 or less than 0.25. These questions are less effective in differentiating users than middling questions.
>
> **Discrimination**
>
> This is the statistical correlation of the question score and the test score, from -1.0 to +1.0. A high correlation (close to +1.0) means that the question is measuring the same thing as the test. A low correlation means that there is little correlation between users getting the question right and users getting a good score in the test. A negative correlation means that getting the question right correlates with getting a bad score for the test. The discrimination is calculated using a standard statistical method (the Pearson product-moment correlation). This method is not reliable if users get negative scores for questions or the test, and you should be careful in interpreting discriminations for such tests. Questions with a discrimination of less than 0.25, or with a negative discrimination have poor correlation with the test score. Users who answer

b. An example of a report for individual learners (one with the lowest mark, and with the highest mark) (from Test 2)

Figure 4: Learner report

User name: L1			**Score**
1	Multiple Response	" Sipho releases a balloon. The b" Answer given : 1 items selected : All balloons move upwards	0/4
2	Multiple Choice	" Which of these gases is most li" Answer given : NH3	0/2
3	Multiple Choice	" Which of the following gases wo" Answer given : CO2	2/2
4	Multiple Choice	"Which of these gases is most li" Answer given: CO2	0/2
5	Multiple Choice	"Which of these gases is most li" Answer given: CO2	2/2
6	Multiple Response	"Select THREE factors, which inf" Answer given: 3 items selected: Temperature, Mass, The strength of electrical current	2/6
7	Multiple Response	"Electrical resistance can be ta" Answer given: 2 items selected: Resistance decreases with length, coiled resistance wires such as the filaments inside a light bulb have more resistance than straight ones of the same length	0/4
8	Multiple Response	"From this list, pick those that" Answer given: 3 items selected: Nucleus, Vacuole, Cell wall	4/6
9	Multiple Response	"Select the only TWO correct sta" Answer given: 2 items selected: A charged object has an equal number of positive and negative charges, Positive charges are stronger than negative charges	0/4
10	Multiple Choice	"The weight of a 12Kg mass on th" Answer given: 12 Newton	0/2
11	Multiple Choice	"A man stands on a bathroom scal" Answer given: The man's weight will decrease	0/2
12	Multiple Choice	"Which of these gases will trun " Answer given: NH3	2/2
13	Multiple Response	"Select THREE statements that ar" Answer given: 3 items selected: They only grow underground, Can give rise to new plants, Can grow on tap roots	4/6
14	Multiple Choice	"The common name of Calcium Hydr" Answer given: Quick lime	0/2
15	Multiple Choice	"The region of the young root of" Answer given : Region of elongation	0/2
			16/48 = 33%

User name	**: L13**	**Score**
1	Multiple Response " Sipho releases a balloon. The b" Answer given: 2 items selected: An upward force is applied on the balloon by the air, All balloons move upwards	2/4
2	Multiple Choice " Which of these gases is most li" Answer given : O2	2/2
3	Multiple Choice " Which of the following gases wo" Answer given : H2	1/2
4	Multiple Choice " Which of these gases is most li" Answer given : H2	2/2
5	Multiple Choice " Which of these gases is most li" Answer given : CO2	2/2
6	Multiple Response " Select THREE factors, which inf" Answer given : 3 items selected: Temperature, Diameter, Type of material	6/6
7	Multiple Response " Electrical resistance can be ta" Answer given : 2 items selected : The wider is the diameter, the less is the resistance, Coiled resistance wires such as the filaments inside light bulb have more resistance than straight ones of the same length	2/4
8	Multiple Response " From this list, pick those that" Answer given : 3 items selected :Nucleus, Mitochondria, Cell wall	4/6
9	Multiple Response " Select the only TWO correct sta" Answer given : 2 item selected : An object can be charged through friction, An object becomes charged when it gains or loses charges	4/4
10	Multiple Choice " The weight of a 12Kg mass on th" Answer given : 120 Newton	2/2
11	Multiple Choice " A man stands on a bathroom scal" Answer given : The pressure **exerted through the one foot on the scale will increase**	2/2
12	Multiple Choice " Which of these gases will trun "Answer given : H2	0/2
13	Multiple Response " Select THREE statements that ar" Answer given : 3 items **selected: They only grow underground, Can give rise to new** plants, Can grow on tap roots	4/6
14	Multiple Choice " The common name of Calcium Hydr" Answer given : Lime **water**	0/2
15	Multiple Choice " The region of the young root of"Answer given : Meristematic	2/2
		35/48 = 73%

There are two important notes to make from the learner report. First, QM Reporter does not show the whole statement of the answers. Therefore, one has to get the statements from the original paper test or read them from QM Designer. Second, the report shows the choices each learner made, the marks each learner obtained for each question, and the total percentage. The report also shows the maximum mark obtainable for each question, and works out the overall percentage.

c. Examples of reports on the whole test

i) <u>An example of a Summary Report</u>

Summary Report 2000/10/25; Test name: College; Number of users: ....
Score: Maximum - 75 %; Minimum - 33 %; Average: 49 %;
Standard deviation: 14.00 % Time to complete test:
Maximum: - 19:55; Minimum: 0:03; Average: 14:50

QM provided information on the average marks, and the standard deviation. The report also shows the times taken to answer the test.

ii. <u>An example of a List report</u>

A list report gives the marks, the name of the test, full names of those who did the test, date on which they did a particular test, and the times each one started and stopped the test. However, the date and times must all be set on the server and computer stations accurately. The times have been excluded from the report because computers were inaccurate with regard to time. QM also kept a record of every attempt each learner made.

List Report 2000/10/25

Score	Test name	User name	Date
33 %	College	L1	2001/10/25
33 %	College	L2	2001/10/25
44 %	College	L7	2001/10/25
50 %	College	L8	2001/10/25
62 %	College	L12	2001/10/25
73 %	College	L13	2001/10/24
75 %	College	L14	2001/10/25

The report indicates the starting time and the stopping time, these did not come through the printout.

QM can arrange the report on demand, according to the marks obtained, name of test, alphabetical order of the learners, date, and time. In the above, we desired a report in order of scores.

Teachers' responses to the interview (See Appendix III)

The responses teachers gave to the interview appear in Appendix III. I have derived a number of themes, which I indicate by an exponent number, and I try to look at these in terms of the sense teachers made of CAA with regard to its diagnostic potential and of the learners' experiences. I have already described the constant comparative method I used to develop these themes (see methods of data analysis).

The sense teachers made of CAA can be summarised into two categories: values and problems, and these can be broken up into 16 themes as shown below.

i. Values (see Appendix III for details)

Theme No. and (Teacher & Answer No.)	**Theme**	**My analysis**
1 (T1- Answer 1)	Interesting, motivating, exciting to learners; encourages learners to do the test	Contributes to motivation
2		
T1- Answer 1	It is helpful – saves time, paper, immediate feedback;	
T2- Answer 3; 11	Convenient (takes burden off the teacher), shows time left; saves time	
3		
T2- Answer 1;	Does not require supervision;	
T1- Answer 3	Learners relaxed	
4 (T2- Answer 1)	Can help with assessing large number of learners, with few teachers	
5		Diagnostic
T2- Answer 1, 8, 11	Can be used to identify learners with problems (so teachers can focus on learners with problems)	
T1- Answer 8)	Can identify problems	
6. (T1 & T2- Answer 2)	Can be used for revision	Diagnostic
7 (T2- Answer 1)	Can show problems, errors, mistakes (so teachers can focus on that chapter)	Diagnostic
8 (T1- Answer 9)	Can show how the learner thinks	Diagnostic

ii. Problems

Theme No.	**Theme**	**My comment**
9 (T2- Answer 9)	Need for skills in diagnosing need how to interpret and to use data	Not only for CAA
10 (T2- Answer 8, 12)	Need for setting appropriate diagnostic questions and administering them	Not only for CAA
11 (T2- Answer 6, 12)	Need for more computers and CAA software Requires a lot money	
12 (T2- Answer 8)	Many questions needed	Not only for CAA
	Need for CAA skills	Requires training

The analysis above indicates that teachers held various values of CAA, including diagnostic values. Further detail of these follows.

a. Potential use of CAA

The potential uses teachers realised are listed as values above. These include motivating learners, saving time, providing immediate feedback, revision, convenience, and offering the possibility of assessing many learners.

It can be seen that they did not consider CAA as a teaching tool, and that teachers did not use the term 'diagnosis' during these interviews, but implied it in saying that CAA can be used to identify learners with problems (so teachers can focus on learners with problems), can be used to identify problems, and that CAA can show how the learner thinks. Teacher 1 (T1) believed that CAA provided data for diagnosis (Question 8), and followed her belief in Question 9 by stating that CAA can be used to show where the problem is, what kind of mistakes learners make, and how the learner thinks. In question 11, Teacher 2 (T2) picked up the point that CAA could help needy students, who he would then pay attention to.

b. Potential problems or inhibiting factors against CAA

It can be noted that all the problems were raised by T2, who was teaching in a more advantaged school. For example, he stated that computers were not enough for CAA, and that CAA and computers were expensive.

Another inhibiting factor was the need for training. In question 8, Teacher 2 identified the need for skills to structure questions for diagnosis with particular reference to what a teacher wanted to achieve from the diagnostic exercise, and the need for many questions.

c. Teachers' opinions about how their learners related to CAA

Responses to questions 1, 3, 5, and 11 relate to what teachers observed when learners were doing the CAT. Responses to question 1 and question 3, indicate that the CAT was fun: i.e., the CAT was interesting, exciting to learners, convenient, and made learners relaxed.

T2 (answer 1) was more impressed by the self-assessment opportunity that the CAT offered to learners, such that the test can '*be done unsupervised*'. He also noted the possibilities it offered such as testing many learners, while enabling attention to '*only those students who don't understand*'.

Learners' opinions after using the CAT – Was Test 2 fair to learners?

A sample of 14 learners' responses to the fairness of Test 2, which they did using computers appear in Appendix IV. This section provides insight regarding the 'face validity of Test 2.

As far as the fairness of the CAT was concerned, the following are themes that learners made (individual learners appear in brackets), derived from Appendix IV:

1. **Good** (L1 - good, L8 - nice)	
2. **Interesting** (L2, L3, & L4)	
3. **Required thinking** (L2)	**This is of diagnostic value**
4. **Difficult** (L1, L3, L6, L8, L11, & L14)	This was expected of diagnostic items
5. **It helps one to revise** (L5, L9)	**This is of diagnostic value**
6. **Enjoyed** (L5) = 1	
7. **Very easy** (L9) = 1	This particular learner obtained low marks!
8. **Encourages study** (L10) = 1	
9. **Okay** (L12) = 1	Fair

Before further analysis, Appendix IV provides data that there was no apparent relationship between 'fairness', 'time a learner took to complete the test', and the 'total marks' a learner earned. Another point that has to be considered is that learners tried to respond in English, which was a 'second language' to all the Blacks, and to some Whites (the Afrikaners). Therefore, their vocabulary was limited, and I had to process some of the responses, either by asking the learner to explain further what s/he meant or by reviewing the statements together with their teachers. Further sources of data were the statements learners made as they did the test. There were instances when I had to use vernacular for the Xhosa learners. In other words, it was necessary to carry out a 'discourse analysis' for some learners' responses.

Learners agreed with their teachers that CAA helped them to revise. Only one learner of the sample believed that the test was fair, as it covered the topics that they had learnt already. After further discussions, the other statements closely related to the test being 'fair' (good, nice, enjoyed, interesting) were found to relate to using the computer. That is, the excitement about using the computer, getting marks immediately, being able to re-do the test, and knowing the time left to do the test. It is not clear though whether it was fair with regard to the way they were taught and assessed.

However, the most popular theme indicated that learners found the CAT difficult (six learners). This judgement is challenged if one looks at the summary report for Test 2, which was made by QM Reporter. These learners obtained an average mark of 49%, in an average time of 14.5 minutes for a test that was set to be done in an hour. With a maximum mark of 75% and a standard deviation of 14%, it is unlikely that their judgement that the test was difficult is correct.

Conclusion

The evaluation should be continuing to include a larger sample. However, for the purposes of obtaining diverse insight into the use of ECPs in South African schools, I deemed it more appropriate to take Zadarh, a totally different ECP from CAA, to schools. The following section is a report on the evaluation of Zadarh.

Introduction

Preliminary studies indicated that teachers told learners facts, and did few practical exercises. It also appeared as though learners were losing interest in biology, and were performing badly at Matric. Investigations indicated that there could be a relationship between enjoyment of biology topics and how easy learners found them. The evaluation of Zadarh explored how it could impact upon teachers' classroom practices and on whether it could arouse learners' interest in biology. This section is a report on the evaluation of Zadarh in Grade 10-12 classrooms.

This study

I borrowed some ideas from the study on CAA, which I had carried out before this one. For example, I approached teachers who were in a Dinaledi project, Carnegie project, or those I was training for an Advanced Certificate in Education' – Science Education, unlike the study on CAA for which I approached schools. In this case I went to those schools whose teachers had agreed to take part. Therefore, these were teachers who could have been enthusiastic about furthering their knowledge.

Secondly, while the evaluation of CAA involved urban-centre schools where learners were used to computers (but not to using them for assessment), the evaluation of Zadarh involved disadvantaged schools in South African townships or rural areas. Most of the teachers and learners had never used computers in their lives. One school where CAA was evaluated was 'advantaged', while the other was not. In evaluating Zadarh, there were only 2 schools that would be considered as 'advantaged'.

Thirdly and fourthly, CAA was based on software designed on a behaviourist paradigm outside SA, while Zadarh was designed on a constructivist paradigm in South Africa, and focussed upon problems that could have been uniquely South African.

Fifthly, I used lessons from the evaluation of CAA to construct an evaluation scheme, which I improved upon in the course of evaluating Zadarh. Therefore, I modified some of the research questions and methods I had used in the evaluation of CAA.

In my opinion, these five differences would yield data that covers a wide scope of experiences of ECPs in South African schools.

Eventually, the study settled on Dinaledi schools because these schools had just been given new computers. The schools were in the Eastern Cape, KwaZulu Natal, and Mpumalanga provinces, and were either in disadvantaged Black townships or rural areas in SA during the third and fourth quarters in 2002 and 2003. The teachers in these schools had very varied qualifications in Biology (see preliminary survey). The numbers of teachers, and therefore learners varied, but dropped to 26 Grade 11 – 12 teachers and 192 learners in 23 schools. Therefore, the evaluation of Zadarh was a much bigger study that the evaluation of CAA, and as such contributes the bulk of the final conclusions and discussions.

As was the case with Adam's (1998) study, this study considered problems in biology classrooms, which validated a need for Zadarh (see preliminary study – Part I).

The difference between this study and Adam's (1998) study is that, while Adams focussed on the effectiveness of Zadarh in learning the relationship between respiration and photosynthesis, using pre-testing and post-testing, and on the use of participants who had computers at their homes, I used Zadarh to evaluate the possibilities that play offers to learning, used qualitative constructivist and interpretative approaches, and chose participants from disadvantaged communities to whom computers were still new in education.

On the basis of the consideration that Zadarh was a constructivist game, the main question for evaluation was: **How does Zadarh help to solve some of the teachers' problems in Grade 10-12 biology classrooms?** The teachers problems in biology classrooms had been investigated in a preliminary survey, but the teachers and I evaluated Zadarh with regard to learners' problems in conceptual understanding of the topics Zadarh deals with (i.e., photosynthesis, and respiration, and evolution), as well as the problems teachers have in their teaching strategies. Subsidiary questions provided answers too this question (Table 15).

Table 15: Research questions and activities on Zadarh

Question		Activity and Tools
1.	**What values do learners and teachers attach to Zadarh?** **What do teachers and learners like/dislike about Zadarh?**	
i.	What in Zadarh makes it desirable / undesirable to learners?	o Observing and interviewing teachers and learners while playing Zadarh, with and without assistance. o Talk-aloud protocol as learners play Zadarh o Video or audio records where participants agreed o Fill-in questionnaires
ii.	How does Zadarh contribute towards understanding selected biology concepts?	
iii.	What stages do learners consider as important when they play Zadarh?	
iv.	How does the interaction with Zadarh relate to the nature of science, and how well does it teach the concepts?	
v.	What outcomes does Zadarh support?	
2.	**How best can Zadarh be integrated in the school?**	
3.	**How best can Zadarh complement the teaching methods employed by the teachers?**	

On addition I evaluated the technical issues relating to Zadarh and school computers. These technical issues included how quickly Zadarh loaded on the computers, and how well it run on the computers.

Methodology

In this case study, I had obtained permission and appointments through the 26 biology teachers who had access to working schools computers. These teachers took letters to their principals I wrote to ask for permission to work in their schools, except for two instances where I faxed the letters.

I demonstrated the installation and operation of Zadarh using my laptop during workshops or using the computers at schools. Teachers then installed and operated Zadarh. I left a Zadarh CD and questionnaire for those teachers and learners. I made a second round of appointments during which I answered teachers' concerns about Zadarh, and then together with the teachers, introduced Zadarh to learners. Teachers promised to give learners opportunities of playing with Zadarh. I made a third round of appointments with the teachers, after they had completed the questionnaire and felt ready for interviews. I outlined what evaluation involves (see manual on evaluation), and reviewed some of the biology concepts with participants during the second and third visit if it was necessary to show teachers and learners how Zadarh should be loaded and started, and if teachers had problems with the evaluation questions. I recorded the dialogues and conversations during the time learners were playing Zadarh. Learners completed a short questionnaire after playing Zadarh.

Findings

Introduction

I have a full file of responses from questionnaires and interviews, which might keep me busy for the next few years. More data continues to come from teachers to the extent that the opinions change as they get used to Zadarh. Therefore, I had to be selective in two ways. The first was to look for statements that provided the information I set to find out, regardless of which question a teacher was responding to. For example, the relevant response to a question on how Zadarh can help learners might appear under a question on how the teacher could use Zadarh in his/her class. The second way was to look for important inputs that inform the quality of user-programme interactions as well as the quality of the evaluation process, which I had not thought about when I was designing the questionnaire and interviews. Quality inputs are not numerical, although the number of people who suggest those inputs matters. I have up to so far confined quantitative data to teachers, and represent samples of statements or themes from learners because of the volume of data from the learners.

In the following section, I provide examples of relevant responses and of important inputs teachers and learners made. I obtained responses from questionnaires applied to teachers, but I had to obtain further clarification by phone, by interviews where I visited the school again, or in workshops organised for CASME work. It should be noted that all teachers claimed that this was their first experience of applying a programme such as Zadarh to solve classroom problems.

Technical aspects relating to using computers

a. Some incidents

A teacher put the mouse on the screen, and another tried to move the 'Zadarh' icon to the 'start' Windows icon. Of course, such incidences are not directly problems of Zadarh, but these represent some of the incidences where it was the first time for teachers and learners to experience computers. Yet, most of these schools had computers for almost a year. Therefore, it was imperative to start by training those teachers who had such problems in basic computer skills.

b. Training

Teachers could run Zadarh without assistance in 3 out of the 23 schools. Training the rest included switching on the computer, logging into the computer, installing Zadarh, using a mouse, and playing Zadarh. Fortunately, all teachers were enthusiastic to use computers and to play Zadarh.

c. Skills / knowledge required for Zadarh

All teachers agreed that Zadarh requires operating a mouse only.

d. Compatibility of Zadarh with school computers

Zadarh was fully compatible and runs well on computers supplied to the Dinaledi schools. Colours were poor and Zadarh was slow on older computers, for example at one school in the Eastern Cape (School 4 EC), even though Zadarh recommended such computer hardware and Windows 3.11.

e. Design and navigation

Teachers found accessing and demonstrating Zadarh to learners easy once it was installed on the computer. Teachers also found that responses of Zadarh to inputs were immediate. One teacher (out of 26 teachers) appreciated the option of restarting Zadarh where you left off, along with your score.

Teachers did not encounter problems in spelling, gender, or cultural biasness, and violence in Zadarh.

Thirty one per cent of teachers found that all icons were easy to click but another 42% complained that the expected actions were not clear. Furthermore, the 'help' option was found useful to only 12 % of the participants. The main complaint was that one found the 'Help' option by trial and error, and that the help provided did not go far enough; i.e., it is limited to showing the cursors.

Other major complaints were about the inability to get out of a situation or bypass some of the puzzles (77%). An example is the room on evolution, where there was no clear way out if one chose to abandon that part. This problem was compounded by lack of a trail one has already taken. All teachers indicated that Zadarh does not show where they had already been, and that there is no guidance on how to move around, especially because some parts look the same.

Using Zadarh to teach

As a teaching resource, one-half of teachers believed that Zadarh was easy to use for teaching the topics it covered, but as a revision exercise. However, teachers raised the following problems prominently:

- The inability to set difficulty levels (58%)
- Ending up at a different point with the same mark. That is, the score is not related to content learners covered. (73%)
- The need for continuous support (i.e., there are no tutorials, loading or running instructions for the programme) (77%)
- There is no way that a learner or teacher can use the programme to find specific information (100%)

Curriculum issues

a. How the programme addresses curriculum issues

Responses to this aspect started by asking each educator what areas of Zadarh they had visited. The majority went as far as respiration and photosynthesis (69%). Only 4% (i.e., one teacher) reached the library where evolution is discussed. All teachers believed that Zadarh was for Grade 12 biology, although 46% thought that the content was not logically arranged.

b. What outcomes does Zadarh support?

Interviews revealed that teachers could not name accurately the specific outcomes for biology, and avoided responding to questions about how Zadarh addresses biology outcomes. Therefore, I had to discuss the outcomes with them in focus groups, when they came to CASME to attend Dinaledi vacation schools.

All of them believed that Zadarh could lead to the achievement of some aspects of the four science specific outcomes, without giving explanations on how. But they also claimed that other, probably cheaper, means such as practical work can be used to achieve these outcomes.

c. Outcomes, which no other method could achieve

Thirty eight per cent said that there are no outcomes from playing Zadarh, which other methods of teaching cannot achieve. Of those (62%) who believed that Zadarh led to enjoyment, 42% thought that this was not important. For example, one argued that Zadarh might *not allow clear teaching except 'playing'.* That is, only 19% thought that enjoyment is a very important aspect of learning, and a game like Zadarh provided it. Nevertheless, teachers also had reservations about playing.

Among the 62%, the main advantage Zadarh offered over practical work was motivation.

However, all teachers were able to note some outcomes:

- Learners enjoyed Zadarh, which meant that they would learn with fun (100%)
- Zadarh required learners to be observant or focussed (46%)

d. Does Zadarh present the nature of science or biology adequately?

This is one of the questions answered by only 12% (3 out of the 26) teachers in the questionnaire. The reason gathered through workshop discussions and follow-up interviews was that most teachers did not comprehend what the Nature of Science (NOS) implied, even after my attempt to provide a simple explanation. The three who responded to this question looked at the NOS in terms of the specific outcomes such that their answers were similar to those they gave in the question about specific outcomes.

What in Zadarh makes it desirable / undesirable to learners? (according to teachers)
(Motivating or repulsive factors)

a. Virtual reality and environment

Sixty nine per cent of teachers thought that only a part of the game simulated real world situations. Complaints about reality include the direction of movement, which is restricted to right angles (100%), and about the speed, which cannot easily be varied (23%). Teachers pointed out exciting realities as the sound of the piano, the sound of closing or opening doors, and the lift (100%).

Similarly, teachers thought that some of the skills used in the game matched those in the real world, especially mental skills such as problem solving (100%). Examples of problem solving the teachers gave include the whole process of extinguishing the fire – i.e., the preparation of carbon dioxide, and include dealing with molecular equations and masses, as well as solving the glucose-pyruvic acid equation.

b. Interactivity and enjoyment

I had to clarify further the question about interactivity. 65% of the teachers believed that the programme was interactive. However, with a statement like '*It is complicated and does not give room for any additional work from the learner or teacher*', this teacher also realised that the player could not introduce new information into the game. They all agreed that the game was very enjoyable – by 'filling the cylinders with oxygen' (92%), and 'extinguishing the fire'. (77%), and scoring after solving a problem (58%).

c. Boring parts

Teachers mentioned getting stuck as the main boring part (62%), which they explained as moving around in the same room', failing to find a solution to a puzzle, or lack of guidance on how to get it right.

d. Suggestions for improvement

Teachers suggested the following improvements

- A manual to state the objectives of Zadarh (81%)
- More help during play, especially when one is stuck (69%)
- Direction – to include diagonal movement (100%)

e. Conclusions on using Zadarh

Teachers would use Zadarh especially to motivate learners (100%), but have the following reservations:

- Time for learners to use Zadarh (100%)
- Computers are often used by computer science (in the three schools offering computer science)
- Zadarh can be used for revision or as an additional source of knowledge.

Learners' evaluation of Zadarh

Data from learners was much more than that from teachers, and was more difficult to analyse because of their poorer English, and volume. We have also to remember that in a game like Zadarh, statistical data might not be meaningful because the Zadarh is an open microworld to the extent that it does not restrict a player in space and time. Learners visited different sites in Zadarh, and so their responses depend upon how many, and which sites they visited. I provide the popular samples of statements to represent the experiences learners had of Zadarh.

It was notable that learners, almost instinctively, consulted each other as they played Zadarh, and in the process talked loudly. So, the talk-aloud protocol that I had planned as a method of collecting data was almost given without request. Secondly, they gathered around one who they believed knew how to go about playing Zadarh, and would go back to their own computers after solving a problem. It was easy to see learners who were either stuck or enjoying an occasion.

Finally, the Play Observation schedule (which is similar to the Science Teaching Observation Schedule – see preliminary survey), was not useful because Zadarh occupied learners throughout, unlike a traditional science classroom.

I have underlined my participation in a sample of classrooms that I managed to attend, and have italicised the learners' statements.

a. Samples of learners' dialogues and notes as they played Zadarh

- About direction

'It goes straight!': They seek assistance from me after a few of them have tried. I show them the help and campus icons.

'There is fire!': They seem stuck, until I start a discussion on how to put the fire out. One learners notices a fire extinguisher – but *'it is out of order'*. *'Where do we get another one?'* 'May be there is one in the house – let us look'
Example of instructions from a colleague: *Straight – Left – Straight – Left – Click – etc.*

- About respiration and photosynthesis

Working out the equation together – one learner acts as a subscriber.
6 molecules x $O_2 + C_6H_{12}O_6 = 6 \times CO_2 + 6H_2O$ *(or* $6 \times H_2O$*). Energy released. Photosynthesis is the opposite of respiration*

- About the safe

School 2 – Eastern Cape
Learners started by making suggestions to one another regarding the meaning of the numbers on the safe. Suggestions included: Molecules; Numbers; Moles; Mass number; Molecular weight, each of which they tried out without success. Looking up to me, I suggested that they should look at the equations of respiration. They worked out the molecular weight for the reactants and then the products, realising that these were equal. But then, this could not sort out entering those numbers on the safe. I had to show them how to enter the number into the safe.

School 2 – KwaZulu Natal
'Safe code = No ? = 6 + 1 + 8 equals 15' 'How do you put the code?' I showed them how to click the safe.
It does not work! It is a molecular weight? OK. Atomic mass of Carbon? And Oxygen and Hydrogen? How many each? 6 x 32 = 192; 6 x 12 = 72 12 x 1 = 12 6 x 16 = 96 'Bring a calculator' OK. Its 372.

- When playing the piano

'Hey, it has real sound!' 'There are molecules on the buttons' Learners work out these codes: *A G D F; F D G A; G F E D C B A;* $CO_2 + H_2O = ?$

- Playing with the blue flower

What is this? Click – OK. Let us write each button and then start again.
Learners clicked randomly on the petals, and were reading the statements that

showed up. They realised that these can form statements about respiration and photosynthesis. They then tried to mentally work them out until I suggested to them to write the statements. So they started writing while noting the position on the leaf as follows: *1st – plant cells; 2nd – animal cells; 3rd – respire; 4th – photosynthesise; 5th – and; 6th – in the light; 7th – dark.* However, after many trials, the learners gave up without results.

b. Notable comments learners made while playing Zadarh

- Whole Grade 12 syllabus should be covered
- More clues needed
- Graphics excellent
- Angles of movement need to be improved

c. Learners' exploratory tendencies

Learners were adventurous much more than the teachers and myself. It was interesting to note that points at which teachers felt stuck or complained for lack of guidance, learners felt that these were challenges or problems to solve, and so in most schools, learners reached places, which the teachers had never seen. Therefore, the enthusiasm and excitement was much more pronounced among learners than among teachers, to the extent that, in all schools, learners requested for copies of Zadarh and asked their teachers to give them a chance to play Zadarh. Obviously, one cannot disregard the possibility that these requests would be ways of gaining access to computers, which schools treat as sacred facilities. In, 19 schools (i.e., 83%), learners had their first chance of entering the computer laboratory with Zadarh.

d. **What stages do learners consider as important when they play Zadarh?**; and **How does Zadarh contribute towards understanding the biology concepts?**

I obtained the answers to these questions from responses learners gave in the questionnaire, immediately after playing Zadarh. See Appendix V for a sample of responses. In Appendix V, the learners are coded: S2-g3 refers to girl number 3, in school number 2. In some cases, one can see the level of thinking of an individual learner by following that learner's responses in the different questions. However, I am still to analyse that dimension.

d i. Events learners remember from playing Zadarh

This section indicates the answers learners gave in the questionnaires to question number 2 (Appendix V): What events do you remember when using this material? We have to be aware that the number of categories or concepts depended upon how many learners reached a particular spot in Zadarh. The answers can be placed in the following three categories.

Category 1 Problematic situations

These are situations in which learners reported that they got stuck.

E.g., *When we got stuck downstairs and didn't know how to get out. Running around in circles (getting lost). Finding one of the doors. Not being able to cross the bridge. When we were looking for a date. I remember where we had shapes and unable to complete them. When we could not get out of the room with a piano.*

Category 2 Problem solving

Problem solving referred to cases where learners realised that they had to do something for them to progress in Zadarh. Most of Zadarh comprised such situations, and so much of the data indicates that learners often had to solve problems.

Examples indicating the role of problem-solving instances include:

- Related to fire

There was fire and we needed to put it out. We solved and managed to put out the fire 'Getting Carbon Dioxide to put out the fire in the store room'. 'I also remember when we struggled to find the door to stop the fire' 'The gas cylinder were a bit tricky because we kind of forgotten the light'. The filling of the tank with carbon dioxide to put out the fire.

- Related to enzymes

Matching the enzymes. The lock and key enzyme action.

- Related to molecular mass

We solved the safe combination of the weight of Carbon, Hydrogen, and Oxygen. When we had to open the safe by calculating the molecular mass. The place where you find glucose and water, and then you have to calculate the *code that you should use as a password. Having to find those two coins with the code. Trying to figure out the coding to open the entrance.*

- Related to the equation of photosynthesis and respiration

We could not solve the piano problem. We spent more than 15 minutes trying to work it out.

Category 3 General activities and knowledge

These were activities, which indicate applications of general biological knowledge.

Getting Oxygen. Getting the gas tank filled with oxygen and distinguishing of the fire in the same room. Trying to refill the oxygen tank. Fire extinguishing and gas tank re-loading. We find the oxygen by first adding carbon dioxide with animal cell then to the pressure. The air cylinder that required air and the photosynthesis experiment. Filling the gas cylinder with carbon dioxide gas by doing respiration on the cells. The collection of carbon dioxide from carbon dioxide from oxygen and cellular respiration. Basics of respiration and photosynthesis. Putting fire out with carbon dioxide. Breathing in oxygen.

d ii. Information learners remember from playing Zadarh

This section indicates the answers learners gave in the questionnaire to question number 3 (Appendix V): What information do you remember from using this material? Their answers can be placed into three categories below.

Category 1 Learners generally remembered the major topics in Zadarh. For example,

- About molecules: E.g., *The fact that you have to calculate add all the molecular weight, The molecular form of compounds, The formula of glucose.*
- About photosynthesis: E.g., *I learnt more about photosynthesis, Oxygen and light is needed for photosynthesis, That under photosynthesis for plants to photosynthesise they need light, water, and carbon dioxide.*
- About respiration: E.g., *respiration requires oxygen, What is needed for cellular respiration, 2 glucose produce 4 pyruvic acid*, and
- About enzymes: E.g., *That only the exact part of the puzzle will fit the enzyme action.*

A few remembered the evolution of man, but this is because fewer learners had visited that topic.

Category 2 Mistaken knowledge

Some learners misunderstood information or actions they took. For example, a statement like '*I must get the oxygen so as to get in the store room*' comes from a learner who misunderstood the need for oxygen in the game. Another statement '*through cellular respiration you can convert carbon dioxide to oxygen or vice versa*' shows that this learner concluded that respiration is reversible. Another learner might think that plants do not respire: '*plants make oxygen from carbon dioxide and animals make carbon dioxide from oxygen*'.

Category 3 Information not part of Zadarh

Some learners provided information that is not obtainable from Zadarh: E.g., Photosynthesis occurs during the day; Humans take in oxygen and breath out carbon dioxide.

d iii. Parts in Zadarh, which learners thought taught them most

This section provides the answers learners made in the questionnaire to question number 4: What part in the process of using the programme teaches you most? I found that their answers fell into the following four categories.

Category 1 Problem-solving

These were situations in where learners felt that they had to do something for them to progress. It seems that problem-solving situations were more informative to learners. Here are examples of statements to that conclusion:

> *Finding the coins to open tools; Searching for the key; The correction of puzzles; Enzymes where we had to put shapes; Calculating molecular masses'; When you try solving the problems of the game; The puzzle-solving; I learnt more when trying to solve a problem especially the piano notes; I think the best was when we tried to gain oxygen from the process of photosynthesis; Finding the combination of the safe – you have to work hard; Using tokens to open something.*

Category 2 Enjoyable situations

These are answers, which indicate that learners were enjoying what they were doing. Playing the piano tops the list of enjoyable events, although it could also belong to category 1 because it was a puzzle. Other learners expressed their satisfaction with the fun involved: E.g., *The manner in which the information is set out. I.e., fun, exciting.*

Category 3 Structures

Learners claim to have learnt from structures such as that of the chloroplast: E.g., *The clear structures of chloroplasts*, and mitochondrion or of the working of the enzymes: E.g., *The working of enzymes; The enzyme pair and finding the code.*

Category 4 General impressions

Some learners were simply impressed by the presentation (E.g., *the chapter of photosynthesis and respiration is well demonstrated than in books*), which in some cases can be linked to the use of virtual reality (E.g., *through interacting and by seeing how things we learn about work actually teaches us a lot; The interaction between the user and the programme*).

d iv. Learners' thoughts on how Zadarh taught them

The question put to learners in the questionnaire was: How does this programme teach you? (Question 5). This question was not very different from question 4 above, but was intended to find out how learners thought they learnt. The answers they provided show that the question was ambiguous. Nonetheless, learners provided some relevant information.

Category 1 Learning from play and fun

Learners felt that playing and fun teaches them well (E.g., *'It is easier to understand when you learn and play at the same time'; 'It teaches me to learn while having fun'; 'It is an excellent teacher, it is fun and it is very organised'; 'You discover a lot of things while playing'; 'It incorporates fun with learning, it isn't a bore or tedious'*), perhaps more than reading textbooks (E.g., *'It is easier to learn than studying a textbook'*).

Category 2 Source of problem-solving skills

There are indications that learners thought they learnt through solving problems, as stated above, and how to solve problems (E.g., *'how to solve a problem'; 'from the enzymes part you learn how to deal with different situation in life'; 'it needs a person to think very hard to figure out new to handle a situation at the same time it teaches you biology'; 'to be attentive and think smartly to keep going'; 'How to basically solve problems with not much information or instructions'; 'it also teaches through puzzles).*

Category 3 Imparts many study skills

Using Zadarh appeared to contribute towards other study skills. Here are examples:

- Perseverance: Examples – *'It shows me that I have to work hard to get something and that I must not stop trying until I get it'; 'by the fact that it is tricky. You have to search a lot'; 'You have to be patient and a great deal of concentration is of essence'.*
- Planning: Examples – *'You need to know exactly where you are heading'; 'It teaches me more on how to plan'.*
- Research or searching for information: Examples – *'It teaches us letting us search and find information for ourselves'; 'It requires you to research and ask why and how'.*
- Memory: ' *and also exercise my memory'; 'It tells/shows you how much you know'; 'To remember things'.*
- Thinking: Examples – *'It challenges the mind'.*
- Observation: Examples – *'It teaches you to actually use your mind and to notice more of your surroundings'; 'That its easy to learn by pictures than reading a textbook and to be very observant inside the rooms'.*

These examples show that Zadarh can inculcate independent study skills.

d v. Advise about Zadarh (also see Table 16)

This was question 6 in the questionnaire: What advice would you give the designers of the programme? (Appendix V). There were undecided learners, who I believe wanted to understand the game more. Almost 50% of learners wanted more fun than facts, and 50% wanted more facts than fun. Then there were those who though that Zadarh was fine the way it was designed. However, the following were some clear advice emanating from the problems that learners experienced with Zadarh.

Category 1 Direction

The prevalent advice concerns direction. Examples towards this view include: *I think it needs change in the direction where you have to go on angle. May be it is better when you want to go to an object go straight to it and when you want to achieve an obstacle, then you can go on angles; Moving towards an object; The movements need to be straight forward (not at 90^0); I'd prefer it or love it more if the whole view of the room was shown at the same time; Shape of the rooms; ... 90^0 direction change to at least 30^0, and; Move in all directions and use buttons to move.*

Category 2 Shape of rooms

By shape of rooms, learners complain about the fact that ground rooms look alike. *E.g., They have to show more of the rooms; The rooms – they are the same, so I would change rooms; The shape of rooms is really confusing; The furniture inside the rooms should be different because they look akin in every room.*

Category 3 Grading the different tasks

I think it should be done in levels and get a particular score to get to the next level; The camera angles and information layout.

Category 4 Zadarh should be easier

Make it easier; few clues; I think more information in the dark room so you have an objective whether you die or not (we want to get anything that would encourage us to score even more'; Every time you get or solve a problem, 'one' should get an instruction on what to do next.

e. Weighing the different aspects of Zadarh

Learners awarded scores in a blank table (see 'A manual for Evaluating Educational Computer Programmes', Part E). The scores in Table 16 are a summary of scores from 60 learners. I picked at random 5 score sheets for girls and 5 score sheets for boys from 6 schools. I held focus-group discussions with each of the five learners to clarify some of the statements. The scores are out of a maximum score of 10.

The scores support the data from questionnaire. Table 16 (arranged from highest to the lowest) shows that learners considered it a better way to learn (score = 9.0), and that Zadarh promotes learning (8.9). They also claimed that they gained more understanding with Zadarh (8.7) possibly because it is fun to use (score = 8.5). In fact, the score recommending more fun (6.9) is slightly higher than that recommending more facts (6.3). This agrees with the learners' comments on *'how Zadarh taught them'*, which indicated that playing adds fun to learning and is desirable.

However, learners thought that the learning aspect of Zadarh was more prominent than the playing aspect (Score = 7.5). A score of 4.8 (the lowest score) confirms this thought further, indicating that the game aspect was less than the learning aspect. They also seem to show that Zadarh should have more content (coverage of content = 5.2), but the content it has is relevant to their syllabus (8.5).

Table 16: Scores learners awarded to various attributes of Zadarh

	Males	**Females**	**Average**
Attribute	**Av. Score/10**	**Av. Score/10**	
It is a better way to learn	8.9	9.0	9.0
It promotes learning	9.0	8.9	8.9
I have gained more understanding than before	9.0	8.5	8.7
Content is relevant to the syllabus	7.4	9.1	8.5
It is fun to use the programme	8.1	8.7	8.5
Feedback	7.7	8.1	8.0
It is too much of learning than a game	7.4	7.5	7.5
Clarity of objectives	6.7	7.3	7.1
I would like more fun	6.4	7.2	6.9
I would like more problems to solve	8.1	5.9	6.7
Organisation of the programme	6.6	6.6	6.6
Objectives are achievable	6.9	6.1	6.4
I would like more facts	7.1	5.9	6.3
Coverage of content	6.8	4.2	5.2
It is too much of game than learning	4.7	4.8	4.8

Overall, males wanted more of the subject matter. That is, male learners wanted more facts or content, and more problems to solve than females, but females scored higher (Table 17). However, females thought that the content was more relevant to their syllabus than males.

Table 17: Differences in scores between male and female learners:

	Males	Females
I would like more problems to solve	8.1	5.9
I would like more facts	7.1	5.9
Coverage of content	6.8	4.2
Content is relevant to the syllabus	7.4	9.1
Average Game score after 1 hour	**188**	**245**

The likely implication of the higher score is that they probably visited more venues in Zadarh and saw that Zadarh covered more content, presented more facts, and presented many problems to solve. Therefore, responses to the four statements above depend upon how far a learner went into the game.

One of the explanations made by some female learners, which I also observed was that males did not easily cooperate with one another – they rarely sought help, and so scored less in tasks that required cooperation. Another observation that might elucidate on why females scored higher is that females made longer answers than boys (Appendix V).

f. Observed or heard complaints

- Zadarh fills up the whole screen, so that one cannot do anything else with the computer without exiting from Zadarh. Therefore, one cannot use other software on the computer without terminating the game
- Zadarh does not load on network stations.
- There is no trail facility to show where you have been

g. Other observations

- No learner ever cared about the percentages of air and carbon dioxide
- Few cared to look at what direction the campus indicated. It appeared as though geographic directions were problematic to learners.

Conclusion

This is rich and diverse data. But it shows important and very deep feelings and experiences individual participants had. Similarities can already be seen between the experiences with CAA and with Zadarh. These come out more clearly in the two discussions that follow.

I start with discussing the findings from the evaluation of CAA, and follow that with the discussion of findings from the evaluation of Zadarh. These discussions will be combined into a single conclusion in the last part of this thesis.

PART VI

DISCUSSIONS

In these discussions we have to bear in mind the contexts in which the teachers and learners experienced CAA and Zadarh, their qualifications and difficulties.

DISCUSSION 1

HOW WELL CAN CAA BE USED TO PROVIDE TEACHERS WITH DIAGNOSTIC INFORMATION ON LEARNERS' CURRENT UNDERSTANDING OF THE SCIENCE CONCEPTS AT GRADE 10 LEVEL?

Introduction

The Joint Committee on Standards for Educational Evaluation (1994) and Fink (1995: 2) describe evaluation as the systematic investigation of the worth or merit of a programme. This same description and others in literature lead discussion 2.

Teachers and learners realised the worth of CAA as claimed by many of the theorists such as Oliver (2000), Gretes & Green (2000), as well as Croft *et al.* (2001). In debating the teachers' opinions of CAA, I need to point out that these teachers (and learners) were fortunately well exposed to computers – that is, there was no apparent threat in using CAA. I particularly focus on those aspects that relate with diagnostic assessment in constructivist environments.

How diagnosing is information provided by CAA?

Linn (2002: 40) states that relevant evidence is necessary for valid interpretations of performance. Here, I focus on whether there was enough data to support diagnosis by trying to answer the following questions: What are the learners' results on the test? What information does CAA provide that is useful for diagnosing learners' knowledge? Moreover; how well do the results indicate the problems that learners have with the topic tested by the CAT?

I need to point out that diagnostic assessment could also be achieved by any other method of assessment, such as a pen-and–paper test. Therefore, in evaluating the diagnostic potential of CAA, it is necessary to differentiate between the qualities of the questions from the contribution to diagnosis that CAA makes. The examiners, who were the teachers in this case, determine the qualities of the questions. The role of CAA is in delivery and processing of marks.

I have had the strenuous experience of analysing results from a diagnostic paper-and-pen test in a Carnegie Project at CASME. I could only manage a sample of 10 scripts, dealing with one question in a week – in fact CASME had to employ somebody to process just over 200 scripts each with five diagnostic questions for a whole year. My experience was similar to that, which Bright (1987:72) describes. I can imagine the stress processing diagnostic data for a class of over 60 learners. The advantage of CAA was promptly analysing large volumes of responses and its speed in providing a feedback. As one teacher pointed out, it was clear that CAA does not diagnose but only provides data immediately to enhance diagnosis.

The analysis of data by QM Reporter helped in identifying recurring mistakes and errors for each individual learner (The New Zealand Council for Educational Research, 2001), and CAA provides detailed data, which teachers can use to diagnose individual learners' problems. The unique data that is normally not available from a paper-and-pen test included standard deviation, facility, and discrimination for each question. This helps in assessing the adequacy of a test item or can be the beginning for identifying problems the whole class has. On addition, QM Reporter indicated the choices that individual learners made, and the number of learners who made a particular choice.

MCQs compared to essays

CAA showed the one advantage MCQs might have over essays: whereas in essays a teacher would have to carry out a 'constant comparative' analysis to identify the most frequent themes or concepts, CAA shows the frequency of choices learners made.

However, justifications in MCQs as alternatives to essays (Tamir, 1996: 98) were possible, but required many alternatives with very fine differences in meaning. CAA can accommodate and analyse a substantial number of alternative answers. In fact Test 3 had one question with up to seven distracters. However, the finer the differences between alternatives are, the more likely that language may play a significant role the way learners interpret the statements. Learners had problems in expressing themselves in English, to the extent that story questions had to be explained many times. Diagnosis and remediation in this case required making sure that learners understood the English vocabulary used in the test. Constructivist dialogues during which learners are encouraged to elucidate upon their conceptions might make diagnosis and remediation more accurate and easier, as well as reduce the language problem.

In my opinion, CAA provided information that is difficult to obtain by other means to enable considerable diagnosis of the learning problems of individual learners.

The capacity of learners to use CAA to diagnose their problems

Learners too realised the usefulness of CAA, mainly because CAA provided learners with opportunities of comparing their performances against earlier ones for each item (ipsative and criterion referencing) by allowing learners to redo the test any number of times they wished. Unlimited redoing of the CAT satisfied another criterion for diagnostic assessment of providing relatively uncontrolled number of tests. They also wanted the CAT to reveal answers (L2 – '*... answers were not revealed at the end. I wanted to see where I went wrong'*). In fact, the option of revealing answers was activated in Test 2. Results indicated that learners improved with practice. This might imply that learners were able to diagnose and remedy their mistakes themselves.

Furthermore, learners were able to evaluate and discuss their answers, with or without assistance from the teachers; thus, self-evaluation was a possibility for individuals or for groups of learners, and at a pace learners desired (Little & Wolf, 1996: xi; Tamir, 1996: 98-99). But these discussions were enabled by the freer atmosphere under which learners did the CAT, and was enforced by the cognitive level at which items were set. So it was much more related to the teaching style than to CAA. Feedback and self-assessment are essential for diagnosis.

Learners enjoyed the CAT because they saw their marks and the correct answers immediately, and were able to redo the test, and again get marked immediately. Their marks improved with time and practice. Anything that can attract learners or motivate learners to test themselves is of great value, and contributes towards self-assessment or self-diagnosis. But this enthusiasm might have been an indication that these learners had been starved of assessment. For example, Teacher 1 had no recent tests or other forms of assessment, and obtained some of the test items for Test 1 from a previous end-of-term examination, and others from a textbook (preliminary survey). Or the enthusiasm could also be due to the rare opportunity for learners to do tests on computers (novelty effect).

Croft *et al.* (2001: 62) as well as Gathy *et al.* (1991: 113) reported positive attitudes towards CAA, and argued that enthusiasm contributed to harder and consistent study, and hence leading to more knowledge and better performance. However, in this case, one teacher associated learner scores more with her own success than with her learners' hard work or learners' difficulties. This association could be beneficial if the teachers would review their teaching strategies with a view to improve, and, although T2 had noted this as a possibility, the two teachers did not comment about their teaching strategies, and did not comment about the data CAA provided.

It is difficult to give a reason for this, but I can speculate that firstly, they saw using CAA in their schools as a remote possibility, and therefore the evaluation exercise as a single event, not worth following up on. Secondly, these teachers were overloaded with many classes and periods. Thirdly, it does not rule out the possibility that they did not know what the next step ought to have been.

The capacity of teachers to diagnose using CAA

Teachers alluded to diagnosis indirectly in their answers. For example, T1 said that data could show both the learner's problems and thoughts (so teachers can focus on learners with problems), and T2 indicated that CAA could show problems, errors, mistakes (so teachers can focus on that chapter). Another diagnostic suggestion teachers brought forth was that CAA can be used for revision and can test large numbers of learners.

Notwithstanding the possibility that these teachers had little time and no reason to attend to this study, the teachers' direct responses did not show that they could diagnose learners' problems. This suspicion is supported by the teachers' apparent reluctance to interrogate the data that CAA produced such as facility, and discrimination. Teachers were worried about marks, but not the kind of responses their learners chose.

For example, Teacher 2 answered that CAA was 'okay' when asked to comment on its diagnostic value. Yet remediation or the re-teaching act after diagnosis requires the teacher to reflect on assessment data and to establish what could have gone wrong in teaching (Bright, 1987: 81). Therefore, the diagnostic value the teachers attached to CAA was compromised. Perhaps Mann (1999) was faced with similar problems in the proposal that other factors should be isolated from such studies, to accurately determine the effects of technology. These teachers had to get more interest or required further training on diagnosis. I tried to give them insight by setting diagnostic items with them. So, I thought that they would have been more enthusiastic about the results and data than the interviews show.

The teachers would have to analyse that information further, for example where distracters are statements involving many concepts. CAA provided the number of learners that have attempted a particular question, and the number of times a 'choice' has been selected, to the extent that a teacher should be able to identify popular misconceptions. The popular misconceptions once identified, can be used in subsequent tests as distracters (Bright, 1987: 78; Tamir, 1996: 97, 107; Maloney, 1987: 510-513; Amir & Tamir, 1994: 94-95; Fraser, 1991: 5-8). A comparison of the frequency of a particular misconception in the first test and in the second test can be used as a measure of the success of diagnosis and remediation.

Diagnosis of learners' problems

Diagnosing learners' problems goes beyond data from QM Reporter. If teachers were to diagnose these learners they not only had to reflect upon their classroom practices, and the reports from QM Reporter, but they would have to look at the responses learners gave during the interview and try to find a pattern or patterns in those responses.

In the interview, a statement such as … *'it was put in a difficult manner so that we can not exactly understand the question'* (L14) can indicate problems with language or a higher level of thinking required in answering a question than what the learner was used to. L4 said that '*needed to know your facts'*, which might mean that the teacher had concentrated on memorising facts (L4). This view is supported by the nature of Test 1, which was set by their teacher but comprised recall items. (Also see class observations in Part I).

That is, teachers would have to make it their practice to discuss tasks constructively with their learners. Indeed, there were instances when learners invited each of us for consultation. This would be in line with Cunningham's (1991: 15) advice on making assessment a naturally arising process, and would be enhanced if the CAT was accompanying practical work. It would even be harder if teachers were not doing practical work (which is most likely). The diagnostic test items, such as Question 6, were contrived from contexts we considered realistic, which required learners to construct plausible interpretations and solutions (Scott *et al.,* 1987: 16; Sagredo as cited in Cunningham, 1991: 16) or from problem situations (Tamir, 1996: 107).

A constructivist would also bring teachers to analysing and helping individual learners (Hannafin, 2004: 13; Pachler & Byrom, 1999: 126, Birenbaum, 1996: 6-7). But, would teachers have the confidence of discussing with their learners as co-learners? I do not know and I doubt on the basis of how teachers conduct their lessons in SA (see Part I & II), and if what I saw in class observations and when learners were doing the CAT is to go by, teachers were shy of critical analysis of questions with their learners.

The way learners dealt with tasks with differentially weighed alternative choices confirms my suspicion that teachers approached science as a collection of facts. Although learners were aware that each choice earned a different score, they treated choices as right or wrong (Scott *et al.,* 1987: 19). The rush through the test by those who obtained low marks implied that learners did not think carefully before choosing answers, showing that processing information was rare to them. This shows a class where things are right or wrong and where giving reasons is rare.

Learners did well in recall items because they had been taught to memorise facts. Hence, the teacher set recall items only. Learners found the test difficult because they had not been assessed at that level of difficulty before. Thus, although on the one hand the Subject Advisor and I sought to improve validity of the tests by covering science processes well, on the other hand, the introduction of higher-level items into the test reduced the validity of that test as far learners were concerned (i.e., the test was difficult), since higher-level thinking and items were not practiced in those classes. The absence of practical work items undermined the validations further. Only recall items were valid in relation to classroom practice, but the tests were invalid in relation to the NOS.

Diagnosing teaching

From the above, it appears that data can provide insight on the teaching style and problems. The results from the tests indicate that there were problems with teaching. I explain below why I state this.

The methods of setting diagnostic tasks as given for example, by Tamir (1996: 97), appeared difficult for teachers, firstly because teachers did not to have the time to think through them, and secondly because teachers seemed to lack the necessary depth of conceptual understanding to set such tasks. The number and plausibility of distracters determined the power of the CAT to diagnose (for example results from Question 6, in Test 2). Each plausible alternative catered for some learners' conception, since each alternative was chosen. However, it is another matter whether teachers would be able to explain the source of such alternative conceptions.

It is possible to identify problem areas using the alternative conceptions learners chose, and to focus upon the source of such problems. Again, referring to data from question 6 (Test 2), 93% of learners believed that the strength of current influences resistance. This belief is logically (rationally) deducible from the relationship $R = \frac{V}{I}$, if one does not do practical work on Ohm's Law. The likely sources of this faulty belief are the teachers since many learners chose this and many common mistakes. Either the teachers did not organise practical work (and taught learners wrong things) or the teachers misinterpreted data from practical work. Pollitt (1990: 885) too observed that the problem is with teaching if many learners make a similar mistake.

Diagnosing, remedying, and learning theories

It follows from my argument above that the errors that learners made in question 6 would be avoided if teachers did practical work. This argument is acceptable to those who believe in empiricism or positivism, as well as constructivist approaches. That is, the argument shows the relevance of science philosophy and learning theories in assessment.

For example, The New Zealand Council for Educational Research (2001) advises teachers to identify the point at which learning faltered. In my opinion, the identification might differ with the learning theory adopted in class. Examiners reports (see Part I) attempted to account for failures in biology; that is, they cannot correct these problems without giving teachers some classroom methodology. However, the dilemma is that classroom practices are rarely located in a single learning theory (see Page 77, Table 10). Thus it might be difficult for a subject advisor or a teacher to come up with a valid diagnosing and remedying strategy. Indeed this might be one of the problems facing South African science education.

Nonetheless, as a starting point, we can, on the basis of experiences such as this evaluation, debate diagnosing and remedying for each learning theory since each recommends a distinct assessment style. The questions in Test 2 can be positioned in behavioural, cognitive or constructivist camps to the extent that the diagnosis and remedy for each question might be different.

For behavioural tasks, Fuchs (1995: 2) advises focussing on discrete domains, and single well-defined concepts or skills. In this regard, items that dealt with single concepts that required memory were valuable. Learners did these well, and the test satisfied Tamir's (1996) and Amir & Tamir's (1994) belief that multiple-choice tests can reveal faulty memories. Behavioural remediation was achieved since learners improved every time they re-did the test - i.e., they eventually identified the correct answers through drill and practice.

A cognitive and constructivist identification and remediation requires a teacher to facilitate practical work and to reason with the learner during the practical and during the time, the learners were solving problems (Hein & Lee, 2000: 7). This did not happen, but we used contrived tasks. Learners were able to discuss with their teachers while they were doing such items. It is unusual for learners to be allowed to talk during a test, and one can argue that these constructivist/cognitive items, which learners considered thought-provoking provided an opportunity to teachers to consider changing their teaching and assessment strategies. That is, the subject advisors would have to see to it that teachers organise practical work, along with assessment. This is apparently the advice in the new C2005, which is supposed to culminate into continuous assessment.

Furthermore, Wood (as cited in Gipps, 1996: 255) advises to identify levels of understanding. I assume that levels of understanding are not behavioural but cognitive or constructivist. Identifying the point of faltering in a constructivist or cognitive sense required multiple choices that represent 'steps' (schemes or constructs) towards conceptual understanding (Hannafin *et al.*, 2004: 13). Such multiple choices can be obtained from learners' alternative conceptions (for example, Test 3 included such alternative conceptions, which learners revealed in Test 2). What I realised is that such questions are challenging to set because they require deep conceptual understanding, a good command of the language they are set in, and an understanding of the learners.

Another difficulty is the assumption that thinking has a road map and is systematic. The way I personally think is certainly not linear, and is haphazard. I would argue that no one could tap in my thinking process and discover accurately where my thoughts are stuck (in the same way, I believe that, one can not teach me how to think but can facilitate experiences that might help my thoughts). Therefore, it seems hard to measure thinking and understanding without resorting to breaking down the processes into smaller, measurable chunks, which seems to be mainly a characteristic of behaviourism, but appears implied in cognitivism and constructivism in as far as schema are constructs arising from interpretations of units of experiences (Cunningham, 1991: 16; Ryan & DeMark, 2002: 67). That is, scientific personal constructs, as perceived by Piaget (as cited in Driver *et al.*, 1994: 6; and in Mwamwenda, 1993: 71) might be assessed by the use of multiple-choice questions that represent different levels of schema construction or conceptions. In this regard, the test again satisfied Tamir's (1996) as well as Amir & Tamir's (1994: 95) advice that multiple-choice tests can be used to reveal the most popular alternative schema.

But this advice brings to the fore Greening's (1998: 23-24) complaint about domesticating constructivism. There are worries that constructivism appears to be more theoretical than pragmatic. 'How well' one understands (Wood as cited in Gipps, 1996: 255) is abstract and interrogates understanding to a level that is so difficult to define, perhaps culminating into researching the concept. South African science classrooms do not indulge into researching, and the levels of difficulty as illustrated by use of rubrics in the new Curriculum 2005 might be hypothetical, as constructivism possibly is. Thus, the attempt to design constructivist MCQ is also an attempt to make constructivist assessment, C2005, and continuous assessment realities. The hypothetical position of constructivism is fortunately a platform for conjectures on possibilities linking MCQs, rubrics, and practical work (continuous assessment), and CAA can support such possibilities (Figure 4).

Diagnosing science processing

Harlen (1993: 28-36) raises the issue of the difficulty of demarcating science processes from each other. This is particularly important in designing rubrics as required by C2005, where it is assumed that each process or outcome can be identified, assessed and remedied. Firstly, the assessment of processes would be easier with practical work, but contrived problems, as stated above, can do. I suggest that we can mainly rely on the final 'compound' outcome, for which we assume that it is a result of a series of science processes.

Figure 5: Rubric, MCQ or Dialogue?

Question on Ohm's Law: What do you know about the relationship between potential difference and resistance?

Level of competency	**Outcomes/constructs/alternative conceptions**	**Practical evidence**	**Score**
			0
1	V is proportional to R	Table of readings // V vs. R	2
2	V proportional to R because charge loses more energy across higher resistance		3
3	V proportional to I if R is constant / the rate at which charge passes R is depends on V if R remains the same // simple V=IR calculation).	Table of readings – draws graph // V vs. I. Determines R	7
4	Factors that affect R affect I with V constant (Factors include, temp, state or phase, thickness, length).	Table of readings // V vs. I under different temperatures - Determines R for each temperature	10
5	Conductors or resistors whose R changes (due to some factors) do not obey this relationship // gives examples	Table of readings and graph // V vs. I. Explains the shape of the line through the points: E.g., for a bulb	13
6	R's in parallel provide more passage to charge and become less effective. Too many Rs allow too many charges to pass thus increasing power consumption or even a short circuit // examples of calculations of current in parallel Rs.	Connects resistors in parallel and records V as well as I.	15

Constructivism can be compatible with MCQs, rubrics, and continuous assessment, but requires a tool such as CAA that can handle multiple measurements and responses. Fortunately, constructivism is also compatible with science (see Part II), and this compatibility is clearer if we take constructivism as a journey, which the brain and body take towards understanding and executing a scientific concept or skill. This compatibility can be simplistically explained. For example, understanding Ohm's Law might include rational or logical deductions (i.e., rationalism), it also might require empirical evidence (i.e., empiricism), and indeed these would be validated either by practical triangulation (i.e., logical empiricism or positivism) or by social negotiations (i.e., social constructivism). In this case a teacher is a facilitator – facilitates triangulations (either logically or socially).

The teacher also facilitates the identification of the stage at which the learner is at to wards constructing a concept, bearing in mind that the journey of constructing is not terminal. The journey should culminate into research, for example, in the form of project work. CAA comes in at this point in that it provides the capacity to the teacher to deal with multiple constructs in form of rubrics (Jonassen *et al.*, 2003: 229), especially in a practical exercise or in a dialogue – it reduces the labour, which, as we know, teachers are complaining about when they deal with continuous assessment and rubrics in classrooms with large numbers of learners.

Thus, MCQs that are constructivist can also serve as a rubric, by which a learner reveals competency by the choice s/he makes. These levels of competency have to be checked upon continuously to see what level each learner is at – hence, continuous assessment, and each level can be given a score.

This model (Figure 4) borrows from Tamir (1996: 96-97); the design shifts to *'correct – best answer'* formats, which contain some *factually correct information.* Besides, no level or construct is wrong (Kuiper, 1996, 1997). The question would be open-ended and present as many alternative constructs as possible.

Problems with CAA

Teachers realised the CAA problems as recorded in Oliver (2000), Taiwo (1995: 19), and Gretes & Green (2000). One of the most crucial problems was that teachers did not possess the appropriate skills for using CAA and of diagnosing learners' problems in science, and that QM was too expensive for the schools. Furthermore, schools could not afford the cost of more computers.

However, these problems do not directly dent the diagnostic value of CAA. Rather, problems challenge choices that schools have to make. The final determinant of such choices is therefore how teachers compare the diagnostic value and other values provided by CAA against other needs of the school. In this case, teachers thought that their schools had other needs that were more valuable than CAA.

Conclusion

This case study showed that CAA is valuable because it instantly provided data teachers and learners could use to diagnose and remedy problems in teaching and learning. This was however inhibited by a shortage of computers and lack of experience to set diagnostic items on a CAA programme. The quality of questions was *inter alia* influenced by the understanding the teachers had of science and of the learning strategies they used. The cost of software for CAA (QM in this case) could militate against the use of CAA in schools. For CAA to be introduced into schools for diagnostic purposes:

- Teachers must be trained about diagnostic assessment, and about using CAA
- The curriculum of the school would have to change to accommodate CAA
- Teachers would need to change their teaching styles, especially to accommodate diagnosis and remedial work.
- There must be a database of common misconceptions on each topic in science
- More computers have to be bought
- CAA companies have to find ways of lowering prices of their products in SA
- Learners should have access to CAA, with guidance from the teacher, any time.

In future, CAA might play a role in constructivist classrooms, by assessing learner's information processing and demonstration of knowledge and skills in virtual laboratories.

The experiences teachers had clearly revealed to them the alternative ways of assessing, diagnosing, and remedying learners' problems. These experiences can change the way teachers run their assessment, and the way they engage their learners.

Lessons from the evaluation of CAA

The process of evaluation indicated that there were benefits from involving teachers and learners in the process. For example, participation and consultation earned the teachers' and learners' cooperation and interest in CAA and diagnostic assessment. Furthermore, the evaluator gets a clearer picture of the schools' capacities in using such programmes, in terms of human and computer resources. I discuss these further under reflections.

However, it will be seen that some of these lessons were beneficial in evaluating Zadarh, which happened after evaluating CAA. The next discussion is about the experiences teachers and learners had of Zadarh.

DISCUSSION 2

HOW DOES ZADARH HELP TO SOLVE SOME OF THE TEACHERS' PROBLEMS IN GRADE 10-12 BIOLOGY CLASSROOMS?

The nature of data

First, the weaknesses of teachers in science and the science curriculum undermined the data. For example, it is evident that teachers lacked knowledge of the Specific Outcomes, and so their critique on whether and how Zadarh achieves these outcomes were compromised. However, this is the reality in some of the schools in SA.

Findings in relation to Adams's and Ivala's evaluation

Unlike Adams' (1998) participants who were probably computer-literate, this evaluation included participants whose competency in computer skills was questionable. Hence, data includes schools' readiness to use computer-based programmes, besides using these programmes in education.

Therefore, the problems Adams recorded about Zadarh, which I also found, are probably insensitive to computer literacy. These include the limitation of movement to forward –backwards, and to left and right, as well as lack of guidance (including a need for a floor plan).

Such problems dispute the conclusion in Adams' (1998: 64) research that playing Zadarh depended on previous experience with the computer. Rather, it appears as though participants improve with experience because that is human nature in whatever they do – i.e., this improvement is not restricted to playing Zadarh or other computer games. The designers have, since Adams' evaluation, attended to the option of starting where one stopped.

This research confirmed the advantages Adams (1998) and Ivala (1998) found from playing Zadarh, albeit in more qualitative detail. Zadarh improved social collaboration among learners, and also improved their understanding of some biology concepts through fantasy, mystery and adequate feedback. I discuss these in more detail below.

Zadarh as a game

Activities in Zadarh fit Pelligrinni's definition of a play (Draper, 2000). Learners enjoyed Zadarh in the sense of fantasy and competed to see who would earn the highest score. Learners also achieved both the *"U-flow" and the "C-flow"* (Rieber, 1996; Draper, 2000) to the extent that some learners played through the school 'break' without realising it. Zadarh managed to instigate 'flow' by drawing learners into deep concentration, providing manageable challenges, providing immediate feedback, and by providing learners with control over their play.

The concern though is striking the right balance between puzzle, important knowledge or skills, and motivation. Zadarh showed that extremely hard puzzles can be de-motivating. For example, learners easily left tasks that appeared hard. On the other hand very easy tasks are eventually unchallenging and uninteresting. For example, there were signs of losing interest in the safe once they knew how it worked. The fact that learners were divided about whether Zadarh should be made more entertaining or more factual bears out this concern and difficulty.

Motivation from Zadarh

One learner summarised the experience with Zadarh as exciting. This excitement is confirmed by many other observations, which can be divided into two categories: The general interest in computers (that individuals generally enjoy manipulating machines which award scores, such as computers) and the discovery that learners can actually solve problems by themselves, especially using a computer.

Zadarh enhanced intrinsic and extrinsic motivations. Intrinsically, learners wanted to see and to know what was around and could explore without guidance. Scores extrinsically enhanced further exploration as well as competition between learners. Another motivating factor in Zadarh is the graphic design and the 3D objects and scenarios as explained by the CIB (2002), as well as (Windschitl *et al.*, 1997). The 3D vision and free movement in a 3D environment excited learners to the extent that they commended the graphics. These observations demonstrated the mutual dependence of extrinsic and intrinsic motivation (Bindra, 1969: 11-12).

Self-regulated learning

The link between motivation and learning (Malone & Lepper, 1987) was indicated by the desire learners expressed to play Zadarh. In fact, learners scored 9/10 for Zadarh as a better way of learning, 8.9/10 that Zadarh promoted learning, and 8.7/10 that they gained more understanding than before. These scores support their answers in part (g), which point at the possibility that Zadarh can inculcate self-directed learning. Learners also realised the benefit Zadarh offers, which is learning with fun, and reported more interest in games than in conventional lessons thus supporting theorists such as Draper (2000), Randel *et al.* (1992: 268), Quinn (1997), and Karaliotus (1999). This conclusion concurs with earlier findings that the difficulty and enjoyment of biology topics are positively and significantly correlated.

Furthermore, Zadarh offers learners the opportunity to regulate the pace – slow learners and those who had to consult textbooks had the time to do that, while knowledgeable ones proceeded at a faster rate. There was also choice – there were learners who looked for particular challenges after browsing through the game. Indeed, there is evidence that Zadarh improved attitudes and motivation towards learning as learners chose what to do and solved problems they chose by themselves (Willis, 2000: 7; Ayayee & Sanders, 1998: 53, 56; Rieber 1996a: 47).

Thus, Zadarh satisfied some of the conditions for self-regulated learning, such as instilling intrinsic motivation and *engaging learners in planning actively.*

Intuition in Zadarh

Perhaps one misunderstood aspect of Zadarh is the difficulty to achieve an optimum balance between enjoyment and the difficulty of puzzles. Learners scored 7.5/10 'too much of learning than a game', which corresponds well with 6.9/10 'I would like more fun'. They also scored 4.8/10 (the lowest score) 'It is too much of game than learning', which is related to 'I would like more facts' (6.3/10). These scores imply that learners wanted more fun in Zadarh, bearing in mind that a balance between fun and learning is needed.

At the same time they wanted more problems to solve - 'I would like more problems to solve' (6.7/10). Hence, it is possible that fun is associated with solving problems. Problems to solve include the suspense due to the player being unaware of where he is, and whether he has been at a site or not. In my opinion, this increases the player's intuition and imagination, and subsequently, higher order thinking skills, which might restructure conceptual understanding (Ross, 1977; Adey, 1987: 19; Wellington, 1994: 24). For example, some of the choices learners made in direction or actions were not logical (Ross, 1977; O'Hear, 1989: 10), but then learners solved more problems than teachers or myself. One particular instance was when a group of learners decided to take a lift although they had not completed previous tasks. The logical thing to do would be to make sure that one completes all tasks and to think that there might be tokens needed when one goes to and into the lift. Other instances of intuition happened when guessing where tokens fitted.

Zadarh achieved most of Medawar's (1969: 55-57) forms of intuition. For example, the puzzle involving fitting together shapes was *inductive intuition*' since these called upon the learners' creativity, while fitting tokens into slots required '*instant apprehension*'. Learners deduced from these activities that each enzyme 'fits' a particular substrate.

Furthermore, Zadarh caters for the "affective" aspect of learning as perceived by Lederman (1998) and Gagné (1985). Zadarh inculcated among players the willingness to collect and use the evidence (respect for evidence) when learners realised for example that they had to calculate the molecular mass of glucose; willingness to change ideas in the light of evidence (flexibility), whenever learners failed to solve a problem and then used some other ideas, and; willingness to review procedures critically (critical reflection) when learners had to refine their playing skills or problem-solving skills.

How Zadarh taught learners

Learners' answers to interviews revealed that they mostly remembered a challenging situation, whether they had solved them or not. It is not far-fetched to assert that learners enjoyed and remembered solving tricky situations, and that they learnt new concepts and skills in the process. These observations support the view that solving problems while playing games contribute towards learning (Randel *et al.*, 1992; Rieber, 1996a: 45; Leutner, 1993: 114; Draper, 2000). Talk-aloud transcripts indicated that learners thought about their thinking (i.e., metacognition occurred) as Birenbaum (1982: 4), Rieber (1996a), and Rieber, *et al.* (1998) suggest. However, they relied upon each other's support or upon guidance from a facilitator, which might imply that they co-constructed not only the concepts but also the processes they had to go through to solve problems.

I have already pointed out a learner who claimed to have understood the structure of a chloroplast better, because Zadarh presents it in three dimensions. This comment supports Windschitl *et al.* (1997) argument that traversing multiple three-dimensional qualitative representations and frames increases conceptual understanding, and leads one to recommend 3D virtual environments.

Zadarh and learning strategies

At the first level, the concern is whether the learning strategy Zadarh uses achieves the objectives for which Zadarh is designed, regardless of whether it is compatible with the teachers' strategies. At another level, can Zadarh fit well enough into teachers' classroom strategies to be integrated into school curricula?

Zadarh employs the three major learning strategies, as recommended in Ertmer & Newby (1993) as well as in Sprinthall & Sprinthall (1990), although it supports constructivism predominantly. I outline aspects of Zadarh that relate to particular learning theories.

Zadarh was behaviourist in associating a correct solution with extrinsic motivation (Fosnot, 1996 8), which was in form of scores, and in encouraging stimulus and response 'cause and effect' relationships (Conway, 1997: 1 – 2; Child, 1997: 10). However, Zadarh does not subscribe to the notion of filling-up learners (Winn, 1997; Child, 1997: 10) because learners use their knowledge and skills to solve problems.

Zadarh has most of the elements of cognitivism such as intrinsic motivation, which was evidenced by learners wanting CDs of Zadarh. Learners had constructed or reconstructed schema from interacting with the social and virtual environments when they claimed to understand concepts better (Piaget as cited in Driver *et al.*, 1994: 5, and Scott *et al.*, 1987: 7). Zadarh improved the learners' '*control and exclusion of variables*' (Adey, 1987: 17-19), when for example; learners prepared oxygen or tried to play a piano. Other well-illustrated aspect of cognitivism included understanding the *frames of reference, and proportion,* since their movement through the game improved with play. These are made apparent in the virtual environment. For example, distant objects are smaller. Zadarh makes learning an adventurous, problem-solving, and discovery activity (Lawton & Hooper, reported in Mwamwenda, 1993: 71; Wollman, 1990: 555).

A test would be necessary to examine the learners' ability to transfer information they learnt from Zadarh to different situations in life (Anderson, *et al.*, 1996). They were however able to transfer knowledge from previous lessons to Zadarh, for example; when they worked out the pyruvic acid stage of respiration. It should be noted that some learners, and teachers too, stated that playing Zadarh was the first activity where they had ever applied knowledge. Whether this kind of applications enhanced the learners' self-concept (Weiner cited in Rieber, 1992: 99) is a matter still to be investigated.

The discussions and arguments between learners as they worked through some problems in Zadarh were reminiscent of cognitive conflict and conceptual change (Tobin & Jakubowski's as cited in Etchberger & Shaw, 1992: 412; West and Pines as cited in Wollman, 1990; Prawat, 1992: page 4, para 3). Zadarh, by the fact that it is a game, provides perturbation or disequilibria as well as awareness of a need to change, when a player finds that s/he cannot solve some problems. For example, learners saw the need to refer to teachers, their textbooks or me when they lost faith in themselves. The learners' commitment to change was evidenced by the decisions they made, which contradicted their earlier understanding. They also realized that the change in understanding required their own introspection and discussions with others (Hannafin & Rieber, 1989: 96). Thus, in all, Zadarh provided platforms for learners' conceptual transformation, especially since new conceptions could be used to solve the problems (Posner, Strike, Hewson, and Gerzog as cited in Wollman, 1990: 555; Geelan, 2000: 4).

Another cognitive aspect Zadarh supports well is spatial cognition as well as higher order thinking skills because learners get a chance of creating meaning by manipulating objects (Osberg, 1997) in the VEs of Zadarh as well as Piaget's stage theory on the ability to comprehend perspective, transformations, ordinal relations, and probability (Cobern, 1996; Patterson & Milakofsky as cited in Osber, 1997). Indeed learners commended Zadarh's visual enhancement of their understanding, better than their textbooks do. Learners showed improvement in their movement in Zadarh, which might imply that their spatial cognition improved. However, this improvement was curtailed by the limitation of movement to right angles only.

The constructivist nature of Zadarh

It is worth noting that learners managed to reach stages that teachers or I could not reach, implying that learners are more inclined to take chances and to explore. The implication is that we teachers limit learners' exploration of learning experiences, and confirms the need for open constructivist environments.

Zadarh is a constructivist game at an individual's level and at a social level, although not all constructs relate to biology concepts. That is, there are activities, which encourage biological constructs, such as calculating the 'molecular mass', but there are also constructs, which are generic or which relate to computer literacy, such as the meaning attached to 3D visual. In both cases, processes such as critical thinking, thinking about one's thinking (metacognition), and problem solving are encouraged, which Yumuk (2002: 142) believes are essential for learning. Jegede (1998: 160) and Yore (2001) argue that such constructivist approaches are essential for learning science.

Playing Zadarh provoked learner's conceptual schemes in some instances (Kuiper, 1994: 280; Shymansky *et al.*, 1997). For example, in calculating the molecular mass of glucose, some learners understood the meaning of the equation better. In other instances, such as fitting the shapes representing the enzymes, Zadarh enhanced constructions of schema (Cobern, 1996: 301; Tamir, 1996: 95) from interpreting experiences and solving problems (Driver *et al.,* 1994: 5; Birenbaum, 1996: 6; Cunningham, 1991: 14; Geelan, 2000; etc).

Zadarh supported autonomy – learners could play alone and make individual decisions and interpretations (Holec, 1979: 3), which Birenbaum (1996: 4) calls interpretative constructivism. However, autonomy included deciding on when a learner could cooperate with others. Hence, unlike findings made by Adams (1998), playing Zadarh also encouraged group discussions during which learners supported each other in a manner Vygotsky (1962) describes as social constructivism.

Multiple intelligences in Zadarh

Of the intelligences listed in McKenzie (2001), the following can be realised by playing Zadarh:

- Visual/Spatial –from illustrations, art, puzzles, anything eye catching, etc.
- Verbal/Linguistic – from discussions while playing Zadarh
- Mathematical/Logical - from reasoning and problem solving
- Bodily/ Kinaesthetic - through movement of the mouse
- Musical/Rhythmic – through playing the piano
- Interpersonal – from discussions
- Naturalist – from biological concepts in Zadarh
- Existentialist – from lessons on evolution

Zadarh and the NOS

A great deal of Zadarh qualifies as rationalist because playing involves rationales based on some prior knowledge and equations or predictions (O'Hear, 1989; Popper, 1974; etc.). Zadarh presents knowledge as facts, but learners challenged this knowledge by their own or by referring to textbooks.

Zadarh combines rationalism with some empirical approaches such as the preparations of oxygen or carbon dioxide from photosynthesis and respiration, respectively (Medawar, 1969: 27, etc.). However, positivism in Zadarh is limited to critical observations (O'Hear, 1989: 14 -19), for example, in terms of knowing where one has been, in the feedback, and in verification of facts. Zadarh also includes spatial frameworks or physical models (Schlick, 1925: 530 – 531) characteristic of objectivity in form of structures of biological concepts such as chloroplasts.

Zadarh satisfies Popper's (1974: 978-984) advice for risky predictions, when a player takes chances during play, although these are not strictly investigative in nature or about science concepts. Learners were able to make what Popper refers to as 'daring hypotheses', but out of intuition. The problem is that risky predictions are of no consequence in Zadarh, other than being stuck and not scoring. Deducting score can possibly make a player more careful, and unfortunately less daring.

But then, Zadarh does not follow any sequence of steps, which scientists follow to answer scientific questions (Lederman, 1998; O'Hear, 1989: 12) – there is no method, since each learner can do anything or can follow a different route. Zadarh relates to scientific inquiry by testing a player's knowledge and by offering opportunities for a player to apply that knowledge (Henry 1975: 62; Wartfosky, 1968: 205; etc.). Furthermore, causality is part of Zadarh as it is with most games, since players start from some prediction, which they go on to try out (Feigl, 1953: 408; Russell, 1929: 390; Wartofsky, 1968; Medawar, 1969; Nagel, 1971). In a game like Zadarh, conditions give some idea to the player what to do.

If we take a laboratory as a place where skills are tried and knowledge is tested, then the virtual environment in Zadarh served as one. A number of NOS processes were achievable. For example, teachers and learners mentioned that they gained manipulative skills, interest in science, group skills, hypothesis testing, finding facts, problem solving, becoming observant, and relating abstractions to reality (Henry, 1975: 61 – 74; Lederman, 1998; DoE, 2002; Tobin *et al.*, 1994: 46). Zadarh also contributed towards scientific inquiry by affording learners opportunities for constructing models (Stratford, 1997: 3-4; Dede *et al.*, 1997; Sanders 2002; etc.). Furthermore, Zadarh provides experiences similar to those experiments offer (Medawar, 1969: 35), such as the preparation of O_2, which can lead to useful observations and eventually to a conceptual framework of generalised expectations (Wartofsky, 1968: 206). Zadarh clearly helps discovery and testing.

Virtual realities or practical work?

It is inevitable for one to compare activities in virtual environments and those in laboratories. It is a general belief that constructions of concepts arise from experiences of phenomena (E.g., in Scott *et al.,* 1987: 7; Harlen, 1993: 28; 2000; etc), and both virtual environments and laboratories offer such experiences, including process-oriented manipulation and transformation of objects (Driver *et al,* 1994: 6; Mwamwenda, 1993: 71; Kumar, 1994: 59), which create disequilibria. Zadarh environments provided the emphases that Yore (2001) recommended for constructivist science classrooms. For example, Zadarh encouraged learners' ideas, discussions and debate, and application of scientific knowledge. Therefore, in my opinion, virtual environments can offer opportunities for scientific inquiry, as do laboratories. There are also instances in Zadarh where perhaps VEs exceed laboratories – for example, preparation of carbon dioxide from plants is not an ordinary activity in school laboratories.

Problems with Zadarh

Playing

One teacher complained that her learners would take Zadarh as a game. However, results do not show any undesirable outcomes that Rieber (1996a) is worried about. I am in agreement with Rieber on the point that outcomes from play might not be revealed through the traditional tests. Zadarh exposed outcomes through dialogue as learners play, and these include enjoyment, and learning of facts.

The other problem was that some problems discouraged learners because they found them very difficult, and there was no support towards solving such problems. As an example, learners ended up at a site surrounded by water. Learners simply left the game, feeling defeated after being perpetually stuck at this site. This observation supports Malone & Lepper's (1987) argument that the level of difficulty and feedback affects motivation and self-esteem.

Weaknesses of Zadarh in terms of the NOS

The first weakness is that designers do not reveal the science philosophy upon which they base Zadarh. Neither did teachers comment on the NOS in the programmes. This does not mean that the designers and teachers did not hold any science philosophy, but it might mean that they did not find it necessary to state those philosophies. This lack of exposition might account for the teachers' and designers' failures to explain why they present science in the way they do. Furthermore, it shows how science instructional designers and educators have ignored the NOS. The NOS should not be ignored in any endeavour to teach science.

Furthermore, the user cannot introduce ideas or experiment on hypotheses. For example, it is not possible to alter variables such as speed in case of Zadarh. Hence, programmes fall into the category of dissipating knowledge as absolute truths (Lederman, 1998). However, these might be limitations of computer technology.

Application of Zadarh in a classroom

Zadarh would fit in afternoon study, after 'normal' lessons, since it might require extended periods. It can be used to introduce concepts it deals with before a lesson. In this case, learners would have to be guided much more. Zadarh can be used as a revision. This is in line with Randel *et al's* (1992: 264) recommendation as pre-course knowledge was found to improve the utility of games. Obviously, learners used pre-Zadarh information to play with more focus. The pre-knowledge is reputed to reduce cognitive load (Wilson, 1995b; Hannafin & Rieber, 1989: 96; Hannafin & Sullivan, 1995).

However, even those who had not studied the concepts fully obtained some understanding, which would make a follow-on lesson easier, and which might motivate learners to find out more. The potential application of Zadarh is undermined by the teachers' attitudes towards play as a possible alternative to classroom instruction.

Conclusion

Zadarh offers much more than the test scores used by Adams (1998) and Ivala (1998) could show, although both researchers reported the learners' excitement and enjoyment as they played Zadarh. The other aspects of Zadarh such as the way it challenges learners with problems to solve are important to enhance interest in and motivation to do biology. Zadarh and other programmes similar to Zadarh, can offer possibilities in dealing with abstract concepts and concepts that are too expensive, lengthy, or dangerous to investigate in a school environment. For all these, I would recommend Zadarh as part of the curriculum regardless of test scores.

There is a lot more to find out and discuss about Zadarh, not least because teachers and learners were changing their opinions with use. This research is endless. The study should continue, and be a source of data for further improvements in ID, in evaluation strategies, and on Zadarh, especially in consideration of the realities in the majority of schools in SA. Zadarh is also important in that it is a South African designed and developed ECP. South Africa needs more of Zadarhs to respond to its local realities.

Play is indeed a source of conceptual frameworks, which should continue through the education systems. Although results did not favour any gender, males seemed to concentrate on non-scoring activities in Zadarh more than females. It might be that the whole problem of fewer females opting for science and related careers starts in their childhood, which restricts them to particular games or particular aspects of games. For example, it might be easier to turn a male into an engineer because the games boys play often involve dismantling and constructing things. These games, together with problem-solving skills, subconsciously continue throughout the life of the male child, to the extent that even those who have not studied engineering end up with some sort of engineering-related hobby such as motor mechanics or building structures. Therefore, building interest towards science-related careers among females ought to start in their childhood, to the extent that we should provide the same toys and games to girls.

In the next part (VII), I analyse the evaluations of CAA and Zadarh with a view of arriving at common benefits and problems.

PART VII

CONCLUSION AND REFLECTION

Introduction

In both evaluations, it is apparent that teachers and learners had to be guided, but had started to learn new information and skills. The most apparent included reviewing their classroom practices, identifying problems, and considered, as alternatives, using ECPs as well as constructivist approaches to learning. Teachers were also exposed to evaluation and to other ways of looking at some science concepts. In this section, I try to explain how these achievements came about.

The quality of these evaluations

Firstly, I have to admit that the evaluation is not complete, and gives the short-term outcomes. However, these qualitative data are richer and can reveal the sophistication and the values of activities than quantitative data. This concurs with the questions on evaluation posed by Heinecke, *et al.* (1999: 1-2), and conclusions in Gredler (2001: 531). The long-term impact caused by the two ECPs will take a longer time to obtain, and would include more summative and quantitative data.

Second, I was much more interested in interpretations, which individual participants made of the two ECPs, the process of evaluation, and how the process affected the way teachers and learners perceived the evaluation process and teaching or learning. I was not seeking a 'final' impact the two ECPs made upon participants. In the interim, the data I have obtained is adequate to reveal the teachers' interpretations.

Third, some of the data would be difficult to quantify in the context of the philosophies and methodology applied, and in consideration of interpretativism in an open learning environment. The way participants interacted with the two ECPs would not be controlled or timed, as this would have contradicted the interpretative, developmental, and constructivist frameworks upon which this evaluation was based. In the case of Zadarh, learners could visit any site, and the case of CAA, learners were free to discuss and ask any questions. In both cases, learners and teachers sought different kinds of support, from each other or from me. Nonetheless, quantitative data can complement the data I have, to give a more complete picture. Further research will extract quantitative data using different philosophical frameworks.

However, this qualitative data is more difficult to analyse, especially in an interpretative paradigm where every individual's statement, activity, and feelings count. I have highlighted those interpretations, which inform the research questions, the interests of the participants, and qualitative aspects of the programmes, which the designers ought to be concerned about basing myself on what they claim to have designed the programmes for. On the other hand, there are simpler issues, which can easily be evaluated quantitatively. These include extent of content coverage, loading a programme, knowledge and skills learners obtain, etc. Checklists and tests can be sufficient for such data.

Evaluations such as these case studies are likely to be undermined by teachers' apparent lack of subject and ID knowledge; the dialogue between the teachers and researchers might not lead to a 'consensus construction' that Lincoln & Denzin (1994: 111) expect, without training. Not surprising therefore, that often, evaluations do not lead to the same conclusions. In the cases researched, the values teachers and learners attach to these programmes are probably not the same as those researched (E.g., the CAA's diagnostic value, and enjoyment of using Zadarh). For example, teachers' worry about financial constraints of their schools as well as their success in teaching as measured by the scores or grades their learners obtain; and not so much about the conceptual framework of the computer programmes. Therefore, it is necessary to make sure that potential end-users understand the programmes, and their aims. Furthermore, teachers should carry out the evaluations so that they can improve their skills in the use of ECPs, in evaluation, and in classroom practice.

On the negative side, these case studies cannot be deployed to become models of operation without further investigation. In that regard, the evaluation process ought to continue with the participating schools, to see how far the teachers develop in their classroom practices, in using ECPs, and in evaluation skills, and with new schools, to see whether the evaluation model works in them.

Reflections on the evaluation process

I reflect on my experiences in the mirror of the research experience I had in agriculture, entomology, and in science education. I am attentive of the changes and differences in research paradigms within and between these fields.

The first observation is that research in ID was more expensive than for example the research I did in science education due to the difficulty in finding schools. Costs were higher because of the time spent on travelling and telephoning schools with well-maintained computer laboratories, and also on training teachers, installing and testing the programmes on the school LAN.

Notwithstanding the fact that any subject can be studied in so many parts of detail, the second observation is that the evaluations comprised a range of investigations, each of which, in my opinion, could stand-alone. These included but are probably not limited to:

- Problems in class and the causes for those problems;
- Professional evaluation on whether the programmes represented the subject they were supposed to teach, and;
- The evaluations of the programmes teachers and learners made on how helpful the programmes were.

As a result, this thesis comprises sections on the NOS, learning theories, a needs analysis in science classrooms (identification of problems), and the evaluations of programmes. These made the thesis lengthy.

These varied knowledge disciplines converged into ID research and showed the dilemma of ID practitioners. Bearing in mind Reeves's (2000a: 20-21) comments that there is sometimes a misunderstanding between basic and applied research based *on* whether research is aimed at extending understanding of the discipline of ID, or whether it was genuinely meant to solve problems, the main dilemma is the possibility that each discipline might interpret the same evaluation data slightly differently. For example, a science teacher might look at how the programme presents the science concepts, totally ignoring the learning theory applied, while a learning theorist would analyse the learning theory used and how it was used in the programme, not bothering about the subject matter. Consequently, I continue to ask myself:

- Is the nature of science and of ID adequately demonstrated through contemporary learning theories? We have to be mindful of the tendency of each of these to impose some order upon nature through laws, which might not be compatible.

- Which of science and learning theories inform the other or is one of them predominant when they interact?
- What should be done if a learning theory compromises the nature of science or of computer use? For example, the act of 'Instruction' is not exactly agreeable with the process of 'Constructivism' in science.

The question of how the nature of science, as well as ID, separately, and then in combination, interact with learning theories, aware that they are ever changing in time and depending upon who defines them is disturbing. Whether learning theories do control disciplines of study, or disciplines of study influence learning theories, are matters that are topical in ID research (Wilson, 1995b). However, the scope of this research does not allow indulgence into these questions. Yet to some extent, I believe that, besides clearly grasping the research methodology and understanding the participants, a clear understanding of the disciplines (E.g., science, psychology, and computer technology), which interact in ID research, and of how they interact, is needed to make proper choices of research methods, to validate research and the computer programmes, as well as to finally come up with a realistic and valid interpretation of the evaluation data. Thus, on the other hand, this dilemma offers opportunities for scrutinising ID, and therefore for a deeper understanding of ID, as well as for changing the way teachers present knowledge. Therefore, evaluation ought to be a compulsory module in ID studies as well as all basic qualifications in education.

Third, the multiple ID perspectives makes subjective interpretations of findings and social constructivism appropriate for ID research at two levels. Interpretativism could be with respect to each discipline involved. For example, interpretativism offers opportunities to scientists to limit their concern to representations of science concepts in Zadarh. At another level, the interpretative paradigm unconsciously, and perhaps against traditional norms of quantitative approaches, allows individuals to express personal views about all the disciplines as played out in a programme without compelling one to a conclusive objective summary that is acceptable to all those disciplines or to other individuals. (I.e., multiple evaluations are acceptable in interpretativism). For example, I am giving my position as the facilitator of the evaluations, and individual schools will give theirs and chose whether and how to use the computer programmes. That is, interpretativism offers schools the freedom of evaluation and choice.

A further challenge therefore, is that participants, each with one's own interpretation, priorities, and sets of values, understood what was going on in some way. Of course, this is true of all experiences in life. Nonetheless, as the facilitator of those evaluations, while still seeking to know and evaluate those interpretations, I see social constructivism playing a role in that all these interpretations are constructs that have to be brought into negotiated values.

Thus, social constructivism is a possible model in evaluation, and brings together the multiple disciplines as well as individual interpretations, while at the same time, it takes care of the need for the participation and interests of stakeholders. The social constructivist model helped us all to exchange ideas from our experiences and disciplines.

COMPREHENSIVE EVALUATION OF COMPUTER EDUCATIONAL PROGRAMMES AS A CHANGE AGENT IN DISADVANTAGED COMMUNITIES

Introduction

This study primarily aimed at designing a comprehensive evaluation scheme for educational computer programmes, and to find out how the teachers' participation impacts upon their classroom practices. There were two main research questions, the first of which was: **What does designing a comprehensive evaluation instrument for educational computer programmes in disadvantaged communities entail?**

Designing a comprehensive evaluation scheme entails understanding the disciplines involved in a programme, and the philosophies behind those disciplines. It also involves getting teachers and learners in the process, which might require training them in the disciplines concerned. I elucidate more on this conclusion in the following part.

Some factors for serious consideration

I have attempted to produce a comprehensive developmental evaluation scheme for

ECPs, and have highlighted issues of concern, which I think Instructional Designers have to consider in the context of schools in developing communities. Data from these studies informs Instructional Designers that the success of computers in such schools will depend upon the understanding teachers and learners have of ID, and of their classroom problems, as well as their ability to evaluate the programmes. Therefore, it is imperative that teachers are trained in this regard, and it is apparent that the evaluation has to be interpretative and constructivist. Teacher constructivist participation implies that they are part of curriculum development and that the evaluation acts as a change agent, which exposes them to possibilities in using computers to improve upon their practices.

In consideration of the above, **what does designing a comprehensive evaluation instrument for educational computer programmes in disadvantaged communities entail?**

1. Understanding the disciplines involved in the programme

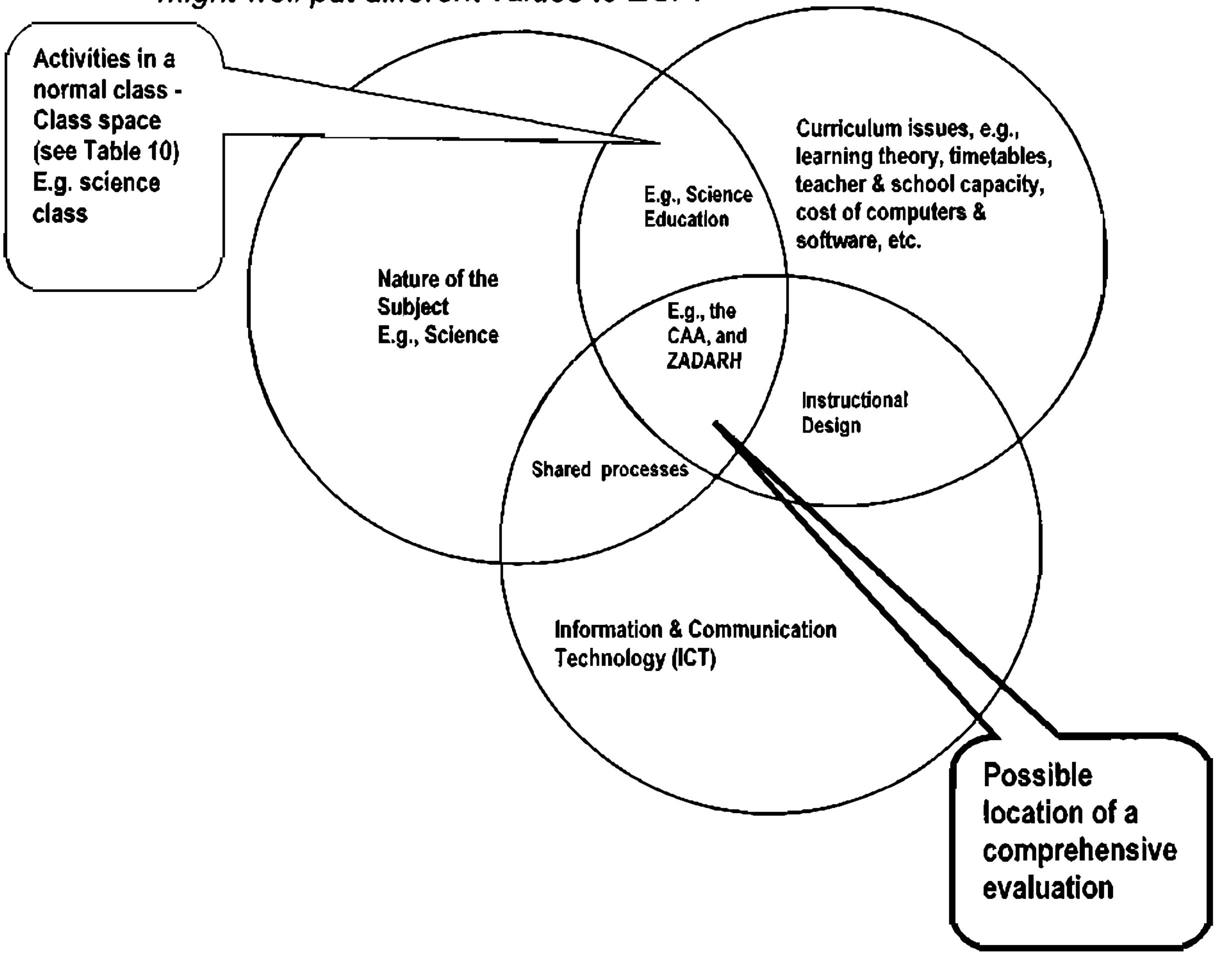
Understanding the disciplines in the programme involves an understanding of the philosophy of the subject which the programme is trying to teach, a consideration of the instructional strategies of the programmes against those preferred by the potential end users or those recommended by the education authorities, and an understanding of how the computer delivers that programme effectively for learners to engage with it (Figure 5).

The implication of adopting the scheme in Figure 6 is that, the teacher-programme interactions have to be designed as interpretative and multidisciplinary in the teachers' view, providing insight into:

i. How well the programme runs and how easy it is to run - technical issues (the hardware and software requirements = ICT),
ii. How the programme addresses curriculum issues (such as learning theories and outcomes), and,
iii. How the programme represents the nature of a subject (the nature of science in this case).

Figure 6 A comprehensive interpretative-constructivist evaluation scheme: 'Evaluate to learn'

The diagram shows the different disciplines that have to be considered in an evaluation – i.e., a scientist, a computer professional, and a teacher might well put different values to ECP.



2. *Involving the teachers and learners in the evaluation process*

First, with reference to Figure 6, the teachers' involvement in the evaluation, hopefully leads to development and action in all the three areas – knowledge of the nature of the subject (improvement in conceptual understanding), skills in ICT or in the use of the computer, and increased understanding of curriculum issues (E.g., constructivism). In this way, the evaluation becomes a change agent, and brings a teacher closer to the instructional designer.

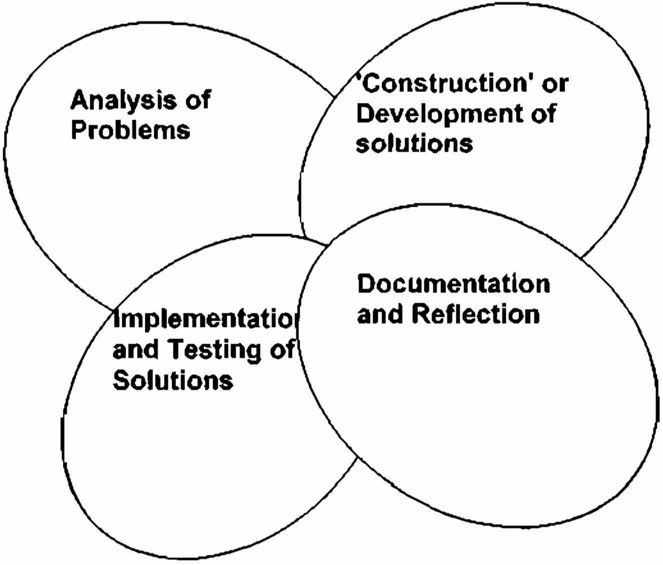
Disparate from other research, evaluations in developing communities simultaneously train, implement the innovation, and the evaluation process. In Africa, the evaluations also involve understanding the cultural paradigms such as *Ubuntu* that operate in schools. This caters for social responsibility in terms of empowering teachers to use computers in schools and in terms of integrating innovations constructively in a manner that is compatible with culture. The approach represent another way that social constructivism is applicable in solving social, economic and development problems. For example, the evaluation model contributes towards improving the teachers' understanding of the curriculum and how the programmes can be integrated in that curriculum, conceptual understanding of the subject the programmes deals with, as well as knowledge of ID.

Second, we have to be mindful of the fact that instructional designers (in an effort to sell their products) campaign to make developing communities accept computer technology as the ultimate solution to many ills in their classrooms. This has created an insatiable desire for anything related to a computer, which might lead to teachers accepting any computer programme, unless the teachers are trained well to carry out the evaluations by themselves.

The evaluation process has to be recursive (see Figure 7) because problems, solutions to those problems, and reflection happen concurrently. Therefore, I propose a different approach to that suggested by Reeves (2000a; 2000b) and Reeves & Hedberg (2003) in two ways that seem to make their model applicable in disadvantaged communities. Firstly, the process of evaluation is recursive because there is no distinct direction (i.e., not necessarily cyclic). It has also to be appreciated that every solution might lead to a new problem or rather that teachers critique more as they learn, and inevitably new needs and questions arise with more experience with the programmes. For example, the 'construction' of solutions was concurrent with analysis of problems – in other words, every solution posed a new challenge or every solution could within it harbour a problem. Take for instance the fact that teachers had to participate constructively in the evaluation process. This however required them to have some basic understanding of evaluation and instructional design, which meant training them during the evaluation.

The second suggestion is that evaluation is not part of a process but evaluation is the whole process. That is, all the elements of the process are evaluated. The teachers ought to evaluate the methods they use to analyse problems, they should evaluate their constructs or solutions, etc.

Figure 7: A developmental and action evaluation model: 'Evaluate to change'



It might not be advisable for experts to evaluate programmes on the teachers' behalf because the teachers are the end implementers who after all have to know whether and how to use the programmes. Rather, the evaluation involves training or organising teachers to do the evaluation. This validates the evaluations, and has the potential to change the teachers' classroom practices. It is possible that many evaluations of computers reveal mixed results because capacities of end-users are also mixed

3. A needs analysis and validation of the programme

Understanding the disciplines involved in the programme and involving the teachers and learners in the evaluation process as stated in the above, lead to a change in their needs. Thus, needs analyses are continuous because needs change. The needs could be in terms of the disciplines illustrated in Figure 5 or in terms of the process illustrated in Figure 6. This step establishes whether a programme contributes towards solving problems at school. However, we need to bear in mind that ID enthusiasts might develop programmes that are not necessarily focussed on a problem, as hardware and software that are more robust become available.

4. The analysis of the capacity of schools to use such a programme

Again related to '1' and '2' above, it is important to establish the capacity of the school to use the programmes. This includes checking on human and computer resources.

5. The influence of a programme upon the teacher's practice

It is important that the way teachers interact with a programme and how it influences the teachers' strategies in the class is noted. Data could comprise the qualitative and quantitative interactions.

As stated in the methodology chapter, these activities (analysis of problems, construction of solutions, documentation and reflection, as well as implementation and testing) happen simultaneously. At the moment though, there is some research in SA that informs the problems and needs, but data has to be refined in detail and breadth to enable designing well-focussed instructional programmes.

These different considerations make an evaluation in developing communities lengthy. However, an evaluator could choose to focus on one of these areas at a time. For example, needs can be established and documented well ahead of producing programmes to address them, and empowering teachers in the use and in the evaluation of computer programmes ought to be part of their training at college. The technical aspects of a programme can also be exhaustively evaluated at the stage of development, especially in light of hardware and software most common in schools. Designers might have to be compelled to state what their programme can and cannot achieve, and to state the philosophies of the subject, as well as of the learning approach they have based their programme upon. These given, the main evaluation remains that of judging how well a programme addresses problems.

Alternatively, it could be that a single programme cannot solve all the areas of concern. For example, CAA might be used in a school for diagnostic assessment, while Zadarh is used for increasing learners' interest in biology. That is, each of the programmes has to be evaluated against a specific use. Designers would then have to state specific areas in the curriculum they are trying to address.

Not least, is the question of the cost and affordability of an ECP, especially compared against other instructional methods. Again, teachers have to know how to cost the other methods (E.g., employing an extra teacher instead of buying an ECP), and to have clear value priorities (E.g., the importance of the quality of assessment offered by CAA, as opposed to buying examination past papers).

I have alluded to social responsibility in the above, but wish to point out that teachers feel over-researched. Every educational innovation requires the teachers' full participation to the extent that they offer genuine and valid data. Therefore, the facilitator of an evaluation has to make sure that teachers gain from the exercise and has to show them how they gain. Teachers were cooperative in these cases because I made them aware of how they benefited from the exercise from the outset. Another way is to register participating teachers for higher qualifications and make these evaluation research projects for that purpose.

OVERALL VIEW OF THE PROGRAMMES

Besides the possible benefits in the above, **how does the teachers' participation in evaluating educational computer programmes influence the teachers' classroom practices?** This second question was unpacked by the following three subsidiary research questions:

a. How well can computer-aided assessment (CAA) be used to provide teachers with diagnostic information on learners' current understanding of the science concepts at grade 10 level?
b. How does Zadarh help to solve some of teachers' problems in grade 10-12 biology classrooms?
c. What values do learners and teachers attach to CAA and to Zadarh?

The discussions in the previous section provide detailed answers to these questions. I reiterate those answers in brief.

CAA provides teachers with data that could be used to diagnose learners' problems, and did this instantly. It also provides evidence on the need for the teacher to be clear of the learning theory s/he is applying in class if diagnosis and remediation is to be carried out successfully.

Computer games can provide better constructivist environments for learning science concepts by granting opportunities for learners to manipulate events and objects, which are not easy to do in real life. As a result, learners enjoy the interactions during playing computer games, especially because computers are tireless. Play makes learning more desirable.

In a nutshell, the constructive participation in the evaluation process provided a deeper understanding of some science concepts, as well as the uses and benefits of the two ECPs to the extent that teachers and learners realised a need to change their teaching and learning strategies. This approach to evaluation is therefore better than that where teachers and learners are simply sources of data.

CHALLENGES

There are common challenges that are revealed from evaluating CAA and Zadarh. The following are the most important in the context of ID in South African science classrooms:

- Improving the teachers' conceptual understanding of science and learning approaches;
- Improving the teachers' abilities to evaluate ECPs;
- Improving the teachers' basic computer skills, and skills to use ECPs;
- Increasing the number of computers in schools so that ECPs can be used more frequently and effectively; and
- Developing ways of integrating ECPs into the school curricula.

These challenges militate against using ECPs in disadvantaged schools.

With regard to evaluation, the main challenge is designing an evaluation that is short but which captures all the relevant information. The other challenge is that an all-inclusive evaluation has to be constructivist and interpretative, especially in the context of developing communities, and both of these philosophies require that the evaluation evolves with the participation of end-users. I.e., ID experts or designers have to become *facilitators*, while teachers become part of the evaluation team. What's more, the interpretations of a programme and values participants' attach to a programme change as they use it. Furthermore, there has to be a follow-up – for example, the evaluation of CAA and of Zadarh is continuing, and so are, hopefully, the improvements of these two ECPs. This scenario leads to an endless revision of evaluation questions and therefore to an endless evaluation, and challenges us with continuously changing the evaluation questions. A comprehensive evaluation scheme is therefore temporary, dynamic, contextual, and, as my experiences show, has to be revised perpetually. Finally, schools need guidelines on using ECPs, which of course have to be revised periodically, in response to new technology, ECPs, and the curriculum.

REFERENCES

Adams, J. C. 1998. *The Use of a Virtual World to Address Misconceptions Held by Students Regarding Photosynthesis and Respiration.* Submitted in partial fulfilment of the requirements for the degree of Master of Science, University of Natal, Durban.

Adey, P. 1987. Science develops logical thinking – doesn't it? Part II. The CASE for Science. *Secondary Science Curriculum Review (SSR), September 1987. (17 – 27).*

Alexander, T. 1997. The Computer as Cognitive Tool: A Constructivist View. [Online] Available: http://www.geocities.com/Athens/Oracle/6002/tool.htm . [13th April 2000].

Amir, R. & Tamir, P. 1994. In-depth Analysis of Misconceptions as a Basis for Developing Research-Based Remedial Instruction: The Case of Photosynthesis. *The American Biology Teacher, Volume 56, No. 2, February 1994.* (94-99).

Amory, A 1997. Integration of Technology into Education: Theory, Technology, Examples and Recommendations. *Contract Report: Open Society Foundation of South Africa (1-21).*

Amory, A. 2001. Visualisation Educational Games. Paper obtained from Prof. Amory (October, 2001).

Anderson, J. R., Reder, L. M. & Simon, N. A. 1996. 'Situated Learning and Education'. *Educational Researcher, 15(4), (5-11).*

Applied Research Laboratory, Penn State University. 1996.
[Online] Available: http://www.umich.edu/~ed626/define.html [19th June 2002].

Atwater, M. .M. 1996. Social Constructivism: Infusion into the Multicultural Science Education Research Agenda. *Journal of Research in Science Teaching. Vol. 33, No. 8, (821-837).*

Ayayee, E. & Sanders, M. 1998. Factors Considered Necessary for Academic Success in First Year Biological Science Courses. *Proceedings of the Sixth Annual Meeting of the South African Association for Research in Mathematics and Science Education. 14-17 January, 1998 UNISA. (52-59).*

Barbera, E. 2004. Quality in Virtual Education Environments. *British Journal of Educational Technology. Vol. 35 No. 1. (13-20).*

Bell, J. 1992. Doing your research project: a guide for first-time researchers in education and social science (First Edition). 325 Chestnut Street, Philadelphia, PA 19106, USA.

Berger, C. F. 1988. How Can Science Teachers and Science Teachers Use Information Technology? In Ellis, J. D. (Ed.) 1989. *Information Technology and Science Education. (73-84).*

Bernt, F. M., Bugbee, Jr., A. C., & Arceo, R. D. 1990. Factors Influencing Student Resistance to Computer Administered Testing. *Journal of Research on Computing in Education.* Vol. 20, No. 3.

Biehler, R. F. and Snowman, J. 1991. *Psychology Applied to Teaching. Seventh Edition.* Boston: Houghton Mifflin Company.

Bindra, D. 1969. The Interrelated Mechanisms of Reinforcement and Motivation, and the Nature of Their Influence on Response. Paper presented at the Nebraska Symposium on Motivation, 1969.

Birenbaum, R. 1982. Games and Simulations in Higher Education. *Simulation & Games. Vol. 13 (1), (3-11) March 1982.*

Birenbaum, M. 1996. Assessment 2000: Towards a Pluralistic Approach to Assessment. In Birenbaum, M. and Dochy, F. J. R. C. 1996 (Eds). *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge.* London: Kluwer Academic Publishers.

Black, H. 1986. Assessment for Learning. In Nuttall, D. L. (Ed.) *Assessing Educational Achievement.* London: The Falmer Press.

Bodner, G. M. 1986. Constructivism: A Theory of Knowledge. *Journal of Chemical Education. Volume 63 Number 10 (873-877) October 1986.*

Bojic, P. 1995. What is Computer-Based Assessment? [Online] Available: http://www.warwic.ac.za/ETS/interactions/Vol2no3/links.htm. [6th October 1999].

Boverie, P., Mulcahy, D. S., & Zondlo, J. A. 1994. Evaluating the Effectiveness of Training Programs. [Online] Available: http://hale.pepperdine.edu/~cscunha/Pages/KIRK.HTM [16th October 2001].

Bright, G. W. 1987. *Microcomputer Applications in the Elementary Classroom. A Guide for Teachers.* Boston: Allyn and Bacon, Inc.

British Educational Communication and Technology Agency (BECTA). August, 2001. Computer Games in Education Project web site. [Online] Available: http://www.becta.org.uk/technology/software/curriculum/computergames/index.html [14th May 2002].

Brotherton, P. N. & Preece, P. F. W. 1996. Teaching Science Process Skills. *International Journal of Science Education, Vol. 18, No. 1. (65-74).*

Bugbee, Jr.,A. C. & Bernt, F. M. 1990. Testing by Computer: Findings in Six Years of Use 1982-1988. *Journal of Research on Computing in Education. Fall 1990: Volume 23 Number 1.*

Burton, J. K., Moore, D. M. & Magliaro, S. G. 2001. Behaviourism and Instructional Technology. In Jonassen, D. H. (Ed.). 2001. *Handbook of Research for Educational Communications and Technology.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. (46-73).

Bybee, R. W. & Ellis, J. D. 1988. A Technology-Oriented Elementary School Science and Health Program: Implications for Teacher Education. *In Ellis, J. D. (Ed.) 1989.* Information Technology and Science Education. (145-162).

Campbell, B. 1998. Realism versus Constructivism: Which is a More Appropriate Theory for Addressing the Nature of Science in Science Education? *Journal of Science Education Vol. 3 - No. 1 - September 1998.* [Online] Available: http://unr.edu/homepage/jcannon/ejse/ejsev3n1.html [21st December 2002].

Castellan, Jr. N. J. 1993. Evaluating information technology in teaching and learning. *Behaviour Research Methods, Instruments & Computers. 25 (2), (233-237).*

Cates, W. M., & Goodling, S. C. 1997. The Relative Effectiveness of Learning Options in Multimedia Computer-Based Fifth-Grade Spelling Instruction. *Education Technology Research and Development. Vol. 45(2) 1997, (27-46).*

Chacko, C. C. 1996. Student Teachers' Views About Difficult and Unfamiliar Topics in Matriculation Biology. *Paper presented at the Annual General meeting of Southern African Association for Research in Mathematics and Science Education (SAARMSE), University of the North, Pietersburg, 24-28 January 1996.*
Child, D. 1997. *Psychology and the Teacher. Sixth Edition.* London: CASSELL.

Christie, P. & Collins, C. 1984. Bantu Education: Apartheid Ideology and Labour Reproduction. In Kalloway, P. (Ed.). *Apartheid and Education: The Education of Black South Africans.* Johannesburg: Ravan Press. (161-183).

Cobern, W. W. 1994. Thinking About Alternative Constructions of Science and Science education. *Proceedings of Second Annual meeting of the Southern African Association for Research in Mathematics and Science Education (SAARMSE), 27-30 January 1994. university of Durban-Westville, Durban, SA. (62-81).*

Cobern, W. W. 1996. Constructivism and non-western science education research. *International Journal of Science Education. Volume 18, Number 3, (295-310).*

Cohen, S., Tsai, F., & Chechile, R. 1995. A model for assessing student interaction with educational software. *Behaviour, Research, Instruments, & Computers. 27(12), (251-256).*

Conway, J. 1997, May. Educational Technology's Effect on Models of Instruction. [Online] Available: http://copland.udel.edu/~jconway/EDST666.htm#cogapp [21st January 2001].

Cohen, L. & Manion, L. 1987. *Research Methods in Education. Second Edition. London: Croom Helm.*

Croft, A. C., Danson, M., Dawson, B. R. & Ward, J. P. 2001. Experiences of Using Computer Assisted Assessment in Engineering Mathematics. *Computers & Education 37 (53-66).*

Crotty, M. 1998. *The Foundations of Social Research.* London. SAGE Publications.

Cunningham, D. J. 1991. Assessing Constructions and Constructing Assessments: A Dialogue. *Educational Technology. May 1991. Volume 31 (5).*

Curriculum Initiatives Branch, State of Victoria (Department of Education & Training), **April 11, 2002.** Teaching and Learning Strategies, [Online] Available: http://www.eduweb.vic.gov.au/sofweb/science/sampleprogram/over/teach.htm [23rd March 2003].

Dalgarno, B. 2001. Interpretations of Constructivism and Consequences for Computer Assisted Learning. *British Journal of Educational Technology, Vol. 32, No. 2, (183-194).*

Dick, W. & Carey, L. 1978. *The systematic design of instruction.*
New York: Scott Foresman and Company.

Dick, W. & Carey, L. 1990. *The Systematic Design of Instruction, Third Edition.*
Glenview, IL: Scott Foresman and Company.

Dede, C. 1995. "The evolution of constructivist learning environments: Immersion in distributed, virtual worlds". *Educational Technology* 35(5) (September-October), (46-52).

Dede, C., Salzman, M., Loftin, R.B., & Ash, K. 1997. Using virtual reality technology to convey abstract concepts. In M. Jacobson & R. Kozma (Eds.), To be published in *Learning the Sciences of the 21st Century: Research, Design, and Implementing Advanced Technology Learning Environments*; edited by Michael J. Jacobson and Robert B. Kozma; Lawrence Erlbaum; late 1997. [Online] Available: http://www.virtual.gmu.edu/pdf/Jacobson.pdf [15th October 2001].

Denscombe, M. 1998. The Good Research Guide for Small-scale Social Research Projects. Buckingham Open University Press.

Denzin, N. K. 1994. The Art and Politics of Interpretation.
In Denzin, N. K. and Lincoln, Y. S. (Eds.) (1994). *Handbook of Qualitative Research.* London: SAGE Publications. (501-514).

Department of Education, South African Government. 23rd December 1998. Vol. 402, No. 19640. No. 6397 No. R. 1718. Government Notice, Department of Education. National Education Policy Act, 1996 (ACT NO. 27 OF 1996). Assessment Policy in the General Education and Training Band, Grades R To 9 and Abet. [Online] Available: http://www.polity.org.za/govdocs/regulations/1998/reg98-1718.html [18th July 2000].

Department of Education, KwaZulu Natal Provincial Examiners' Reports, 1997; 2001; 2003.

Department of Education, South African Government Annual Report 1998. [Online] Available: http://www.edufound.org.za/infor.htm [22nd January 1999].

Department of Education, South African Government. 31st May 2000. Report of C2005 Review Committee. Executive Summary. [Online] Available: http://education.pwv.gov.za/polici...E%20SUMMARY%20OF%20CURR%2020005.html [7th September 2001].

Department of Education, South African Government. 2002. Revised National Curriculum Statement Grades. [Online] Available: http://education.pwv.gov.za/DoE_Sites/Curriculum/Final%20curriculum/policy/English/Natural%20Sciences-print.pdf [16th December 2003].

Doll, W. E. Jr. 1989. Foundations for a Post-Modern Curriculum. *Journal of Curriculum Studies.* 21(3), (243–253).

Draper, S. W. 2000. Analysing fun as a candidate software requirement. [Online] Available: http://staff.psy.gla.ac.uk/~steve/fun.html#2 [15th October 2001].

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. 1994. Constructing Scientific Knowledge in the Classroom. *Educational Researcher. October 1994*, (5-12).

Duffy, T. M. & Cunningham, D. J. 2001. Constructivism: Implications for the Design and Delivery of Instruction. In *Handbook of Research for Educational Communications and Technology.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. (170-198).

Duhem, P. 1953. Physical Theory and Experiment. In Feigl, H. & Brodbeck, M. (Eds.). 1953. *Readings in the Philosophy of Science.* New York. Appleton-Century-Crofts, Inc.

Einstein, A. 1940. The Fundaments of Theoretical Physics. In Feigl, H. & Brodbeck, M. (Eds.). 1953. *Readings in the Philosophy of Science.* New York. Appleton-Century-Crofts, Inc.

Eisenhart, M. A. & Howe, K. R. 1992. Validity in Educational Research. In LeCompte, M. D., Millroy, W. L., and Preissle, J. (Eds.). *The Handbook of Qualitative Research in Education.* New York: Academic Press Inc.

Elliot, J. 1991. Changing contexts for Educational Evaluation: The challenge for methodology. In Lewy, A. (Eds.) 1991. *Studies in Educational Evaluation. Oxford. Pergamon Press, (215-238).*

Enslin, P. 1984. The Role of Fundamental Pedagogics in the Formulation of Educational Policy in South Africa. In Kalloway, P. (Ed.). *Apartheid and Education: The Education of Black South Africans.* Johannesburg: Ravan Press. (138-147).

Etchberger, M. L.. & Shaw, K. L. (1992). Teacher Change as a Progression of Transitional Images: A Chronology of a Developing Constructivist Teacher. *Teacher Change.* Vol. 92(8), December.

Ertmer, P. A. & Newby, T. J. 1993. Behaviorism, cognitivism, constructivism: Comparing critical features from in instructional design perspective. *Performance Improvement Quarterly, 6 (4), (50-70).*

Feigl, H. 1953. Notes on Causality. In Feigl, H. & Brodbeck, M. (Eds.). 1953. *Readings in the Philosophy of Science.* New York. Appleton-Century-Crofts, Inc.

Fink, A. 1995. *Evaluation for Education and Psychology.* London: SAGE Publications.

Fontana, A. & Frey, J. H. 1994. Interviewing. In Denzin, N. K. and Lincoln, Y. S. (Eds.) (1994). *Handbook of Qualitative Research.* London: SAGE Publications.

Fosnot, C. T. 1996. Constructivism: A Psychological Theory of Learning. In *Constructivism: Theory, Perspectives, and Practice.* Teachers' College, Columbia University. Teachers, College Press, New York.

Fraser, W. J. 1991. The Logic of Educational Evaluation. In Dreckmeyr, M. & Fraser, W. J. (Eds.). *Classroom Testing in Biology and Physical Science.* Bloemfontein: HAUM Tertiary.

Fuchs, L. S., 1995. Connecting Performance Assessment to Instruction: A Comparison of Behavioral Assessment, Mastery Learning, Curriculum-Based Measurement, and Performance Assessment. [Online] Available: http://ericec.org/digests/e530.htm [2001, March 23].

Gagne, R. M. 1985. *The Conditions of Learning and Theory of Instruction.* (4th Ed.) NY: CBS College Publishing.

Gardner, H. 1993. *Multiple Intelligences. The Theory in Practice.* Harper Collins, New York.

Gathy, P., Denef, J-F., & Haumont, S. 1991. Computer-Assisted Self-Assessment (CASA) in Histology. *Computers Education. Vol. 17, No 2.* (109-116).

Gardner, P. L. 1975. Science and the Structure of Knowledge. In Gardner, P. L. (Ed.) 1975. *The Structure of Science Education.* Hawthorn Victoria. Longman.

Gay, L. R. & Airasian, P. 2000. *Educational Research. Competencies for Analysis and Application. Sixth Edition.* Columbus, Ohio: Merrill.

Geelan, D. R. 2000. Sketching Some Post-modern Alternatives: Beyond Paradigms and Research Programs as Referents for Science Education. [Online] Available: Electronic Journal of Science Education - V5 N1 - http://unr.edu/homepage/crowther/ejse/geelan.html [20th July 2000].

Gibbons, A. S. 2000. The Practice of Instructional Technology. *Paper presented at the AECT 2000, Annual International Conference of the Association for Educational Communications and Technology, Denver, CO, October, 2000.* Online. [Available] http://www.aect.org/Intranet/Publications/Technology_101.pdf [30th October 2002].

Gipps, C. 1996. Assessment for Learning. In Little, A. & Wolf, A. (Eds.). *Assessment in Transition. Learning, Monitoring, and Selection in International Perspective.* Oxford. Pergamon.

Gredler, M. E. 2001. Educational Games and Simulations: A Technology in Search of a (Research) Paradigm. In Jonassen, D. H. (Ed.). 2001. *Handbook of Research for Educational Communications and Technology.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. (521-540).

Greene, J. C. 1994. Qualitative Program Evaluation. Practice and Promise. In Denzin, N.K. and Lincoln, Y. S. (Eds.) (1994). *Handbook of Qualitative Research.* London: SAGE Publications, (531-542).

Greening, T. 1998. Building the Constructivist Toolbox: An Exploration of Cognitive Technologies. *Educational Technology / March-April 1998, (23-35).*

Gretes, J. A. & Green, M. 2000. Improving Undergraduate Learning with Computer-Assisted Assessment. *Journal of Research on Computing in Education.* Volume 33. Number 1. Fall 2000.

Guba, E. G. & Lincoln, Y. S. 1989. *Fourth Generation Evaluation.* Newbury Park, CA: Sage.

Guba, E. G. & Lincoln, Y. S. 1994. Competing Paradigms in Qualitative Research. In Denzin, N. K. and Lincoln, Y. S. (Eds.) . *Handbook of Qualitative Research,* (428-444) London: SAGE Publications.

Haladyna, T. M. 2002. Supporting Documentation: Assuring More Valid Test Score Interpretation and Uses. In Tindal, G. & Haladyna, T. (Eds.). 2002. *Large-Scale Assessment Programs For All Students. Validity, Technical Adequacy, and Implementation.* Mahwah, New Jersey. Lawrence Erlbaum Associates, Publishers. (Pages 89-108).

Hannafin, M. J. 1999. Learning in Open-Ended Environment: Tools and Technologies for the Next Millenium. [Online] Available: http://it.coe.uga.edu/itforum/paper34/paper34.html [30th October 2002].

Hannafin, M.J., Hall, C., Land, S., & Hill, J. 1994. Learning in open environments: Assumptions, methods, and implications. *Educational Technology, 34*(8), (48-55).

Hannafin, M. J., & Hannafin, K. M. 1991. The Status and Future of Research in Instructional Design and Technology Revisited. In G. Anglin (Ed.), *Readings in Instructional Technology (302-309).* Littleton, CO: Libraries Unlimited.

Hannafin, M. J., Hannafin, K. M., Hooper, S. R., Rieber, L.P., & Kini, A. 1996. Research on and Research With Emerging Technologies. In D. Jonassen (Ed.), *Handbook of Research in Educational Communications and Technology* (378-402). New York: Macmillan.

Hannafin, M. J., Hannafin, K. M., Land, S. M., & Oliver, K. 1997. Grounded Practice and the Design of Constructivist Learning Environments. *Educational Technology Research and Development, 45(3). (101-117).*

Hannafin, M. J., Kim, M. C., & Kim, H. 2004. Reconciling Research, theory, and Practice in Web-Based Teaching and Learning: the Case for Grounded Design. *Journal of Computing in Higher Education. Spring 2004, Vol. 15(2), (30-49).*

Hannafin, M. J., Reeves, T. & Hayden, J. J. 2001. Understanding and Addressing Stakeholder Interests: Evaluation and the World of Policymakers. In W. Heineke & J. Willis (Eds.), *Research Methods for Educational Technology* (251-267).

Hannafin, M. J. & Rieber, L. P. 1989. Psychological Foundations of Instructional Design for Emerging Computer-Based Instructional Technologies: Part I. *Educational Technology, Research and Development. 37 (2), (91-101).*

Hannafin, R. D. & Sullivan, H. J. 1995. Learner control in Full and Lean CAI Programs. *Educational Technology Research and Development. Vol. 43. No. 1 1995, (19-30).*

Hargis, C. H. 1995. Second Edition. Curriculum Based Assessment. A Primer. Illinois. Charles C. Thomas.

Harlen, W. 1980. Evaluation in Education. In Straughan, R. & Wrigley, J. (Eds.) 1980. *Values and Evaluation in Education.* London. Harper & Row Publishers.

Harlen, W. 1991. Pupil Assessment in Science at the Primary Level. *Studies in Educational Evaluation. Vol. 17, (323-340).*

Harlen, W. 1993. *The Teaching of Science.* London. BPCC Wheaton Ltd..

Harlen, W. 2000. Assessment in the Inquiry Classroom. [Online] Available: http://www.nsf.gov/pubs/2000/nsf99148/lcd/ch_11.htm [25th April 2001].

Hart, E. P. & Robottom, I. M. 1990. The science-technology-society movement in science education: A critique of the reform process'. *Journal of Research in Science Education, 17(6), (575-588).*

Herrington, J. 2002. Designing authentic activities for Web-based courses. In M. Driscoll & T. C. Reeves (Eds.), *Proceedings of E-Learn 2002* (Montreal, Canada). Charlottesville, VA: Association for the Advancement of Computing in Education.

Hein, G. E. and Lee, S. 2000. Assessment of Science Inquiry. [Online] Available: http://www.nsf.gov/pubs/2000/nsf99148/lcd/ch_12.htm [25th February 2000].

Heinecke, W. F., Blasi, L., Milman, N. & Washington, L. 1999. New Directions in the Evaluation of the Effectiveness of Educational Technology. Paper given at Papergiven at *The Secretary's Conference on Educational Technology-1999.* [Online] Available: http://www.ed.gov/Technology/TechConf/1999/whitepapers/paper8.html [30th October 2002].

Hendricks, A. J. 2002. Investigating whether using the computer as a tutor could enhance 1st year Peninsula Technikon science students conceptual understanding of electrical concepts. *Proceedings of the 10th Annual Association for Research in Mathematics, Science and Technology Education. 22-26 January 2002. University of Natal, Durban KwaZulu-Natal,* (II 47-53).

Henderson, J. & Wellington, J. 1998. Lowering the language barrier in learning and teaching science. *School Science Review, March 1998. 79(288),* (35-46).

Henry, N. W. 1975. Objectives for Laboratory Work. In Gardner, P. L. (Ed.) 1975. *The Structure of Science Education.* Hawthorn Victoria. Longman.

Heron, J. 1996. *Co-operative inquiry. Research into the Human Condition.* London: SAGE Publications.

Hickey, D. T., Horwitz, P., D. T., Kindfield, A. C. H., & Christie, M. A. T. 2003. Integrating Curriculum, Instruction, Assessment, and Evaluation in a technology-Supported Genetics Learning Environment. *American Educational Research Journal. Summer 2003, Vol. 40, No. 2, (495-538).*

Hickey, D. T., & Zuicker, S. J. 2002. A New Perspective for Evaluating Innovative Science Programmes. *Science Education. 2002*. (539-563).

Hitchcock, G. & Hughes, D. 1995. *Research and the teacher: A qualitative Introduction to School-Based Research. 2nd Edition.* London: Routledge.

Hogle, J. G. 1996. Considering Games as Cognitive Tools: In Search of Effective 'Edutainment'. University of Georgia Department of Instructional Technology. (ERIC Document Reproduction Service No. ED 425737).

Holloway, R. E. 2001. Diffusion and Adoption of Educational technology: A Critique of Research Design. In Jonassen, D. H. (Ed.). 2001. *Handbook of Research for Educational Communications and Technology.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. (1107-1133).

Horn, R. E. & Cleaves, A. 1980. *The Guide to Simulation/Games for Education and Training. 4th Edition.* Sage Publications: Los Angeles.

Huberman, A. M. & Miles, M. B. 1994. Data Management and Analysis Methods. In Denzin, N. K. and Lincoln, Y. S. (Eds.)(1994). *Handbook of Qualitative Research.* (428-444) London: SAGE Publications.

Huysamen, G. K. 1994. *Methodology for the Social and Behavioural Sciences.* Southern Book Publishers.

Imenda, S N & Muyangwa, M M. 1996. *Introduction to Research in Education and Behavioural Sciences.* Umtata: University of Transkei.

Ivala, E. N. 1998. *Identification of Misconceptions Held by Teachers and Students with Respect to Concepts of Mendelian Genetics and Assessment of Teaching Methods to Overcome Such Misconceptions.* Submitted in partial fulfilment of the requirements for the degree of Masters of Education, University of Natal, Durban.

Jacobs, G. 1992. Hypermedia and discovery-based learning: a historical perspective. *British Journal of Educational Technology. Volume 23, Number 2. 1992, (113-121).*

James, R. K. 1988. Using Change Theory to Manage the Implementation of Educational technology in Science Classrooms. In Ellis, J. D. (Ed.). 1989. *Information Technology and Science Education. (Pages 187-205).*

Jegede, O. 1998. The knowledge base for learning in science and technology education. In Naidoo, P. & Savage, M. (Eds.) *African Science and Technology Education into the New Millennium: Practice, Policy and Priorities, (151-176).*

Jenkins, E. W. 1999. Practical work in School Science some questions to be answered. In Leach, J. & Paulsen, A. C. (Eds.) 1999. *Practical Work in Science Education. Recent Research Studies.* Rosenoerns, Frederiksberg C, Denmark. Roskilde University Press.

Jennings R., & Everett, D. 1996. Education for servitude? A survey of “out-of-school youth” in South Africa: Eastern Cape. *Designed and analysed for the out—of-school Children and Youth Policy and Research Initiative and the Department of Education by the Community Agency for Social Inquiry (CASE).*

Jonassen, D. 1991. Thinking technology. *Educational Technology, 34(4), 34-37.*

Jonassen, D. H., Howland, J. L., Moore, J. L., & Marra, R. M. 2003. *Learning to Solve Problems with Technology. A Constructivist Perspective. Second Edition.* Upper Saddle River: Merrill Prentice Hall.

Karaliotus, Y. 1999. The Element of Play in Learning - The Role of Synergetic Playful Environments in the Implementation of Open and Distance Learning Online [Available]: http://users.otenet.gr/~kar1125/proj99.htm. [2nd October 2001].

Kasalu, L., Doidge, M., & Sanders, M. 2002. Evaluating a web-based package on human population issues for teachers. *Proceedings of the 10th Annual Association for Research in Mathematics, Science and Technology Education. 22-26 January 2002. University of Natal, Durban KwaZulu-Natal, (III 160-165).*

Kiboss, J. 1998. Relative effects of a computer-based instruction in physics on students' attitude, motivation and understanding about measurement and perceptions of classroom environment. Unpublished PhD thesis, University of the Western Cape.

King, M. & Van den Berg, O. 1992. *Success or Failure? Examinations and Assessment. Focus Series Two.* Pietermaritzburg: Centaur House.

Kneale, W. 1949. Induction, Explanation, and Transcendent Hypotheses. In Feigl, H. & Brodbeck, M. (Eds.) 1953. *Readings in the Philosophy of Science.* New York. Appleton-Century-Crofts, Inc.

Knight, M. & Brown, A. 2000. Computer based Assessment. [Online] Available: http://ctiweb.cf.ac.uk/HABITAT/HABITAT4/compass.html [14th December 2001].

Kozma, R. 2000. Reflections on the State of Educational Technology Research and Development. *Educational Technology Research and Development. Vol. 48, No.1 2000, (5-15).*

Kraft, D., & Sakofs, M. (Eds.). 1988. The theory of experiential education. Boulder, CO: Association for Experiential Education.

Kuhn, T. S. 1974. Logic of Discovery or Psychology of Research. In Schilpp, P. A. (Ed.). *The Philosophy of Karl Popper Volume 14 (Part II)* Le Salle, Illinois. Open Court.

Kuiper, J. 1994. Student ideas of science concepts: alternative frameworks? *International Journal of Science Education. Vol. 16, No 3, (279-292).*

Kuiper, J. 1996. Educating scientists for the future. *Higher Education, June 1996.*
Kuiper, J. 1997. Quirks and *Quarks*: Changing Paradigms in Educational Research. *Meeting of the Association for Research in Mathematics, Science and Technology Education. 22-26 January 1997. University of Witwatersrand, Johannesburg. (530-534).*

Kuiper, J. 1998. Science and Culture: The Art of Balancing. This appears as a paper in SAARMSE, but was handed out by Prof. Kuiper during a MEd. Course at Rhodes University, 1999.

Kumar, D. D. 1994. Hypermedia: a Tool for Alternative Assessment? *Educational Technology. Volume 31. Number 1.*

Laurillard, D. 2000. How can the non-narrative media support the learner's own narrative construction? [Online] Available: http://meno.open.ac.uk/meno/
[30th October 2002].

Lawton, D. & Gordon, P. 1996. *Dictionary of Education.* Second edition. London: Hodder & Stoughton.

LeCompte, M. D., Preissle, J., with Tesch, R. 1993. *Ethnography and Qualitative Design in Educational Research. Second Edition.* San Diego: Academic Press.
Leutner, D. 1993. Guided Discovery Learning with Computer-Based Simulation games: Effects of Adaptive and non-Adaptive Instructional Support. Learning and Instruction Vol. 3, (113-132).

Lederman, N. G. 1998. The State of Science Education: Subject Matter Without Context. [Online] Available: Electronic Journal of Science Education V3 N2 - December, 1998 - ... , and also
http://unr.edu/homepage/jcannon/ejse/lederman.html [16th July 2002].

Levine, D. K., 2001. Economic and Game Theory. What is Game Theory? Online. [Available]: http//levine.ssnet.ucla.edu/general/whatis.htm [26th November 2004].

Lincoln, Y. S. & Guba, E. G. 1985. *Naturalistic Inquiry.* London. SAGE Publications.

Lincoln, Y. S. & Denzin, N. K. 1994. The Fifth Moment. In N. K.Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research. 118-137.* Thousand Oaks: SAGE Publications Inc.

Linn, M. C. 1988. Science Education and the Challenge of Technology. In Ellis, J. D. (Ed.) 1989. *Information Technology and Science Education, (119-144).*

Linn, R. L. 2002. Validation of the Uses and Interpretations of Results of State Assessment and Accountability Systems. In Tindal, G. & Haladyna, T. (Eds.). 2002. *Large-Scale Assessment Programs For All Students. Validity, Technical Adequacy, and Implementation.* Mahwah, New Jersey. Lawrence Erlbaum Associates, Publishers, (27-66).

Linnerman, S. R., Lynch, P., Kurup, R. & Bantwini, B. 2002. South African science teachers' perceptions of the nature of science (NOS): towards a framework for curriculum discussion. *Proceedings of the 10th Annual Association for Research in Mathematics, Science and Technology Education. 22-26 January 2002. University of Natal, Durban KwaZulu-Natal. Pages III 205-210.*

Little, A. & Wolf, A. (Eds.) 1996. *Assessment in Transition. Learning, Monitoring, and Selection in International Perspective.* Oxford. Pergamon.

Lloyd-Jones, R., Bray, E., Johnson, G., and Currie, R. (Eds.) 1986. Assessment. From Principles to Action. London: MacMillan.

MacDonald, C. J. & Farres, L. 2003. Constructivist Instructional Development Models: A Tool for Examining Context Diversity. DRAFT. [Online] Available: http://www.cade-aced2003.ca/conference_proceedings/MacDonald.pdf
[29th November 2004].

MacDonald, M. A. & Rogan, J. M. 1988. Innovation in South African Science Education (Part I): *Science teaching observed. Science Education 72(2), (225-236).*

Mackay, L. D. 1975. The Role of Measurement and Evaluation in Science Courses. In Gardner, P. L. (Ed.) 1975. *The Structure of Science Education*. Hawthorn Victoria. Longman.

Madaus, G. F. & Kellaghan, T. 1992. Curriculum Evaluation and Assessment. In Jackson, P. W. (Ed.) 1992. *Handbook of Research on Curriculum*. (119-154). New York. Macmillan Publishing Company.

Makhurane, P. & Kahn, M.1998. The role of science and technology in development. In Naidoo, P. & Savage, M. (Eds.) *African Science and Technology Education into the New Millennium: Practice, Policy and Priorities, (23-33).*

Malone, T.W. & Lepper, M.R. 1987. Making Learning Fun: A taxonomy of Intrinsic Motivations for Learning. *In Snow, R.E. and Farr, M.J. (eds.) Aptitude, Learning and Instruction III: Cognitive and Affective Process Analysis. Erlbaum, Hillsdale, N.J.*

Maloney, D. P. 1987. Ranking tasks. A New Type of Test Item. *Journal of College Science Teaching*, May 1987, (510 – 515).

Manzini, S. 200. Learners' Attitudes Towards the Teaching of Indigenous African Science as Part of the School Science Curriculum. *Journal of the Southern African Association for Research in Mathematics, Technology and Science Education. Volume 4, Number 1, (19-32).*

Margenau, H. 1974. On Popper's Philosophy of Science. In Schilpp, P. A. (Ed.). *The Philosophy of Karl Popper Volume 14 (Part II)* Le Salle, Illinois. Open Court.

Martin, L. M. W., Hawkins, J., Gibbon, S.Y. & McCarthy, R. 1988. Integrating Information Technologies into Instruction: The Voyage of the Mimi. In Ellis, J. D. (Ed.) *1989. Information Technology and Science Education, (173-186).*

Matthews, M. R. 2000. Appraising Constructivism in Science and Mathematics Education. In Phillips, D. C. (Ed.) 2000. *Constructivism in Education. Opinions and Second Opinions on Controversial Issues. Ninety-ninth Yearbook of the National Society for the Study of Education. Part I.* Chicago, The University of Chicago Press.

Maykut, P. & Morehouse, R. 1994. Beginning qualitative research. A philosophical and practical guide. London: The Falmer Press.

McComas, W., Clough, M., & Almazroa, H. 1998. The Role and Character of the Nature of Science in Science Education, (3-39). In McComas, W. F. (ed.) 1998. The Nature of Science in Science Education. Rationales and Strategies. Science & Technology Education Library. London. Kluwer Academic Publishers.

McKenney, S. & Van den Akker, J. 2002. Computer-Based Support for Science Education Materials Developers. *Proceedings of the 10th Annual Association for Research in Mathematics, Science and Technology Education. 22-26 January 2002. University of Natal, Durban KwaZulu-Natal, (III 406-417).*

McKenzie, W. 2001. *It's Not How Smart You Are -It's How You Are Smart!.* Howard Gardner's Theory of Multiple Intelligences. Online [Available] http://surfaquarium.com/mi.htm [16th July 2002].

McLellan, H. 2001. Virtual Realities. In Jonassen, D. H. (Ed.). 2001. *Handbook of Research for Educational Communications and Technology.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. (457-487).

Medawar, P. B. 1969. *Induction and Intuition in Scientific Thought.* London. Methuen & Co. Ltd.

Mergel, B. 1998. Instructional Design & Learning Theory. [Online] Available: http://www.usask.ca/education/coursework/802papers/mergel/brenda.htm#Behaviorism [19th June 2002].

Merriam, S. B. 1998. Qualitative research and case study applications in education. Revised and expanded from case study research. San Francisco: Jossey-Bass.

Mgujulwa, J. & Kenyon, A. 1994. Attracted to Science: A Close Look at a Primary Science Lesson. *Proceedings of Second Annual meeting of the Southern African Association for Research in Mathematics and Science Education (SAARMSE), 27-30 January 1994. University of Durban-Westville, Durban, SA, (260-268).*

Miller, L. & Olson, J. 1994. Putting the computer in its place: A study of teaching with technology. *Journal of Curriculum Studies, 26 (2): p121-141].*

Millward, L. J. 1998. Focus Groups. In Breakwell, G. M., Hammond, S., & Fife-Schaw, C. (Eds.) 1998. *Research Methods in Psychology.* London. SAGE Publications Ltd.

Moshell, J. M., Hughes, C. E., & Loftin, B. 1999. Virtual Reality as a Tool For Academic Learning. Draft. Online [Available] http://vehand.engr.ucf.edu/handbook/Chapters/chapter52.PDF [15th October 2001]

Mosimege, M. D. 1997. The Use of Games in Mathematics Classrooms. *Proceedings of the Fifth Annual Meeting of the Association for Research in Mathematics, Science and Technology Education. 22-26 January 1997. University of Witwatersrand, Johannesburg, (530-534).*

Moje, E. B. 1995. Talking about Science: An Interpretation of the Effects of Teacher Talk in a High School Science Classroom. *Journal of Research in Science Teaching. Vol. 32, No. 4, (349-371).*

Mphahlele, M. K. 1996. Supervision of science education research: critique of the discourse. Proceedings of the Fourth Annual meeting, 25 to 28 January. South African Association for Research in Mathematics and Science Education. 236-249.

Munro, R. G. 1975. Curriculum Evaluation. In Gardner, P. L. (Ed.) 1975. *The Structure of Science Education.* Hawthorn Victoria. Longman.

Muraskin, L. D. 1993. *Understanding Evaluation: The Way to Better Prevention Programs.* U.S. Department of Education, contract number LC89089001, task order number LC900940.

Muwanga-Zake, J. W. F. 1998. Research Portfolio. Submitted as a partial fulfilment of the requirements for the award of a degree of Master of Education (Science Education).

Muwanga-Zake, J. W. F. 2000. Is Science Education in South Africa in a crisis? The Eastern Cape Experience. *Journal of the Southern African association for Research in Mathematics, Technology and Science Education. Vol. 4 (1), (1-11)*
Mwamwenda, T. S. 1993. *Educational Psychology. An African Perspective.* Durban: Butterworth Publishers (Pty) Ltd.

Myers, M. D. 2000. Qualitative Research in Information Systems. [Online] Available: wysiwyg://3/http://www.Auckland.ac.nz/miss/isworld/index.htm [2000, May 21].

Nagel, F. 1951. The Causal Character of Modern Physical Theory. In Feigl, H. & Brodbeck, M. (Eds.). 1953. *Readings in the Philosophy of Science.* New York. Appleton-Century-Crofts, Inc.

Nagel, E. 1953. Teleological Explanation and Teleological Systems. In Feigl, H. & Brodbeck, M. (Eds.). 1953. *Readings in the Philosophy of Science.* New York. Appleton-Century-Crofts, Inc.

Nagel, E. 1971. *The Structure of Science.* London. Routledge.

Naidoo, P. 2002. Subject Adviser, Biology. KwaZulu Natal Provincial department of Education. Personal interview. [19th September 2002].

Natal College of Education, Piermaritzburg. 1997. *Teacher Education: Routes to Competency. The Teacher in the Classroom. An Introduction to Educational Theory and Practice for South African Students.*

National Center for Research on Evaluation, Standards, and Student Testing (CRESST) **6th Oct. 99.** [Online] Available: http://www.cresst96.cse.ucla.edu/CRESST/CRESSTLAUSD/glossary.htm [5th June 2000].

National Council of Teachers of Mathematics, Republic of South Africa, May 1995. *Assessment Standards for School Mathematics. Continuous Assessment.* Unpublished Notes.

National Research Foundation, Republic of South Africa. 2004. [Online] Available: http://nrfonline.nrf.ac.za/templates/Guide%20to%20Research%20Support%20for%202005-The%20NRF%20Focus%20Area%20Programme.doc [6th May 2004].

Neuman, W. L. 1997. Social Research Methods. Qualitative and Quantitative Approaches. Third Edition. London. Allyn and Bacon.

New Zealand Council for Educational Research. 2001. [Online] Available: http://arb.nzcer.org.nz/nzcer3/nzcer.htm [21th February 2002].

Nichols, R. G. & Allen-Brown, V. 2001. Critical Theory and Educational Technology. *Handbook of Research for Educational Communications and Technology.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. (226-252).

Noffke, S. E. 1995. Action Research and Democratic Schooling. Problematics and Potentials. In Noffke, S. E. & Stevenson, R. B. (Eds.) 1995. *Educational Action Research. Becoming Practically Critical.* New York: Teachers College Press. (1-12).

Ogunniyi, M. B. 1996. Science, technology and mathematics: the problem of developing critical human capital in Africa. *International Journal of Science Education. Vol. 18, No. 3, (267-284).*

Ogunniyi, M. B. 1997. Multiculturalism and Science Education Research in the New South Africa. *Proceedings of the Fifth Meeting of the Southern African Association for Research in Mathematics and Science Education. 22-26 January 1997. University of Witwatersrand, Johannesburg, (50-53).*

Ogunniyi, M. B. 2000. Teachers' and Pupils' Scientific and Indigenous Knowledge of Natural Phenomena. *Journal of the Southern African Association for Research in Mathematics, Technology and Science Education. Volume 4, Number 1, (70-77).*

O'Hear, A. 1989. *An Introduction to the Philosophy of Science.* Bungay, Suffolk. Richard Clay Ltd.

Oliver, A. 2000. Computer Aided Assessment – the Pros and Cons. [Online] Available: http://www.Herts.ac.uk/ltdu/learning/caa_procon.htm
[14th December 2000].

Olsen, J. B. 1990. The Four Generations of Computerized Testing: Toward Increased Use of AI and Expert Systems. *Educational Technology/March 1990, (36-41).*

Olsen, J. B., Maynes, D. D., Slawson, D, & Ho, K. 1989. Comparisons of Paper-Administered, Computer-Administered and Computerized Adaptive Achievement Tests. *Journal of Educational Computing Research, Vol. 5(3), (311-326).*

Ormell, C. 1980. Values in Education. In Straughan, R. & Wrigley, J. (Eds.) 1980. *Values and Evaluation in Education.* London. Harper & Row Publishers.

Osberg, K. 1997. Spatial Cognition in the Virtual Environment. [Online] Available: http://www.hitl.washington.edu/publications/r-97-18/ [29th December 2000].

Osberg, K. M., Winn, W., Rose, H., Hollander, A., Hoffman, H. & Char, P. 1997. The Effect of Having Grade Seven Students Construct Virtual Environments on their Comprehension of Science. [Online] Available:
http://www.hitl.washington.edu/publications/r-97-19/ [29th December, 2000].

Pachler, N. & Byrom, K. 1999. Assessment of and through ICT.
In Leask, M. and Pachler, N. (1999). (Eds.) *Learning to Teach Using ICT in the Secondary School.* London: Routledge.

Pajares, F. 1998. The Structure of Scientific Revolutions. [Online] Available: http://www.emory.edu/EDUCATION/mfp/Kuhn.html [21st October, 200].

Papert, S. 1993. The children's machine: Rethinking school in the age of the computer, New York, NY: Basic Books.

Pausch, L. M. & Popp, M. P. 1997. Assessment of Information Literacy: Lessons from the Higher Education Assessment Movement. [Online] Available: http://www.ala.org/ala/acrlbucket/nashville1997pap/pauschpopp.htm
[28th September 2000].

Pecorino, P. A. 2001. *A First Course In Philosophy. An Online Text and Course.* Available:
http://www2.sunysuffolk.edu/pecorip/SCCCWEB/ETEXTS/INTRO_TEXT/default.htm
[17 th February, 2003].

Peled, Z., Peled, E. & Alexander, G. 1991. Ecology and Experimentation in the Evaluation of Information Technology Interventions in Natural Classroom Settings. In Lewy, A. (Eds.) 1991. *Studies in educational Evaluation. Volume 17 number 2/3. Oxford. Pergamon Press, (419 – 448).*

Percival, F. & Ellington, H. 1984. *A Handbook of Educational Technology.* London: Kogan Page.

Perkins, D.N. 1991. Technology meets constructivism: Do they make a marriage? *Educational Technology, 31* (5), (18-23).

Perkins, D. N. 1996. Preface: Minds in the 'Hood. In B. G. Wilson (Ed.), *Constructivist learning environments: Case studies in instructional design.* Englewood Cliffs NJ: Educational Technology Publication.

Phillips, D. C. (Ed.) 2000. An Opinionated Account of the Constructivist Landscape. In *Constructivism in Education. Opinions and Second Opinions on Controversial Issues. Ninety-ninth Yearbook of the National Society for the Study of Education. Part I.* Chicago, The University of Chicago Press.

Pinkus, J. 1996. *Foucault.* [Online] Available: http://www.massey.ac.nz/~ALock/theory/foucault.htm [30th October 2002].

Pitman, M. A. & Maxwell, J. A. 1992. Qualitative Approaches to Evaluation: Models and Methods. In LeCompte, M. D., Millroy, W. L., and Preissle, J. (Eds.). *The Handbook of Qualitative Research in Education.* New York: Academic Press Inc.

Pollitt, A. 1990. Diagnostic Assessment Through Item Banking. In Entwistle, N. (Ed.) 1990. *Handbook of Educational ideas and Practices.* London and New York: Routledge.

Popper, K. 1974. Replies to My Critics. In Schilpp, P. A. (Ed.). *The Philosophy of Karl Popper Volume 14 (Part II)* Le Salle, Illinois. Open Court.

Prawat, R. S. 1992. Teacher's Beliefs about Teaching and Learning: A Constructivist Perspective. *The University of Chicago.*

Question Mark. [Online] Available: http://www.qmark.com/ [9th September 1999].

Quinn, C. N. 1997. Engaging Learning. [Online] Available: http://itech1.coe.uga.edu/itforum/paper18/paper18.html [15th October 2001].

Ramsey, G. 1975. Science Teaching as an Instructional System. In Gardner, P. L. (Ed.) 1975. *The Structure of Science Education.* Hawthorn Victoria. Longman.

Randel, J. M., Morris, B. A., Wetzel, C. D., & Whitehill, B. V. 1992. The Effectiveness of Games for Educational Purposes: A Review of Recent Research. *Simulation & gaming, Vol. 23 No. 3, September 1992, (261-276).*

Reeves, T. C. 1994. Systematic Evaluation Procedures for Interactive Multimedia for education and Training. In Reisman, S. (Ed.). *Multimedia Computing: Preparing for the 21st Century.* Harrisburg, PA: Idea Group. ***Also*** [Online] Available : http://mime1.marc.gatech.edu/MM_Tools/MMDM.html [20th September 1999].

Reeves, T. 2000a. Socially Responsible Educational Technology Research. *Educational Technology/November-December. (19-28).*

Reeves, T. 2000b. Enhancing the worth of instructional technology research through 'design experiments' and other development research strategies.
Paper presented at the Annual AERA Meeting, April 24-28, New Orleans).

Reeves, L. P. & Hedberg, J. G. 2003. Interactive Learning Systems Evaluation. Englewood Cliffs, New Jersey 07630.

Rieber, L. P. 1992. Computer-Based Microworlds: A Bridge Between Constructivism and Direct Instruction. *Educational Technology, Research and Development. 40 (1), (93-106).*

Rieber, L. P. 1996a. Seriously Considering Play: Designing Interactive Learning Environments Based on the Blending of Microworlds, Simulations, and Games. *Educational Technology Research & Development. Volume 44. Number 2* (43-58).

Rieber, L. P. 1996b. Animation as Feedback in a Computer-Based Simulation: Representation Matters. *Educational Technology Research & Development. Volume 44. Number 2. No. 1. 1996. (23-42).*

Rieber, L. P., Smith, L., & Noah, D. 1998. The value of serious play. *Educational Technology, 38(6), (29-37).* [Also online] Available: http://itech1.coe.uga.edu/~lrieber/valueofplay.html [30th October 2002].

Richey, R. & Nelson, W. 1996. Developmental research. In D. Jonassen (Ed.), *Handbook of Research on Educational Communications technology, (1213-1245).*

Robottom, I & Hart, P. 1993. Towards a meta-search agenda in science and environmental education. *International Journal of Science Education. Vol. 15, No. 5, (591-605).*

Rollnick, M. 1998. Relevance in science and technology education. In Naidoo, P. & Savage, M. (Eds.) *African Science and Technology Education into the New Millennium: Practice, Policy and Priorities, (79-90).*

Rosenblueth, A. 1970. *Mind and Brain. A Philosophy of Science.* Cambridge. The MIT Press.

Ross, K. L. 1977. The Arch of Aristotelian Logic. The Doctrine of the Prior and Posterior Analytics. [Online] Available: http://www.friesian.com/arch.htm [15th July 2002].

Ross, K. L. 1998. The Beginning of Modern Science. [Online] Available: http://www.friesian.com/hist-2.htm [20th October 2001].

Ross, K. L. 1999. Sir Karl Popper (1902-1994). [Online] Available: http://www.friesian.com/arch.htm [15th July 2002].

Russell, B. 1929. On the Notion of Cause, with Applications to the Free-Will Problem. In Feigl, H. & Brodbeck, M. (Eds.). 1953. *Readings in the Philosophy of Science.* New York. Appleton Century-Crofts, Inc.

Ryan, J. M. & DeMark, S. 2002. Variation in Achievement Scores Related to Gender, Item Format, and Content Area Tested. In Tindal, G. & Haladyna, T. (Eds.). 2002. *Large-Scale Assessment Programs For All Students. Validity, Technical Adequacy, and Implementation.* Mahwah, New Jersey. Lawrence Erlbaum Associates, Publishers. (Pages 67-88).

Salvia, J. & Ysseldyke, J. E. 1988. *Assessment in Special and Remedial Education.* Fourth Edition. Boston: Houghton Mifflin Company.

Sandals, L. H. 1992. An Overview of the Uses of Computer-Based Assessment and Diagnosis. *Canadian Journal of Educational Communication, Vol. 21, No. 1 (67-78).*

Sanders, M. 1995. A useful Format for Questionnaires and Tests Used in Educational Research. *Proceedings Vol.2 SAARMSE Third Annual meeting. January 1995, Cape Town, (712-725).*

Sanders, M. 2002. Secondary school biology learners' difficulties in interpreting diagrams of biological sections. *Proceedings of the 10th Annual Association for Research in Mathematics, Science and Technology Education. 22-26 January 2002. University of Natal, Durban KwaZulu-Natal. II 85-90.*

Sanders, M. & Khanyane, M. 2002. The interpretation of biology textbook illustrations by Grade 10 learners. *Proceedings of the 10th Annual Association for Research in Mathematics, Science and Technology Education. 22-26 January 2002. University of Natal, Durban KwaZulu-Natal, (III 364-370).*

Sanders, M. & Mogodi, G. 1998. Terminology Problems Experienced by Standard Eight Ecology Pupils. *Proceedings of the Sixth Annual Meeting of the South African Association for Research in Mathematics and Science Education. 14-17 January, 1998 UNISA, (314-319).*

Sanders, M. & Mokuku, T. 1994. How Valid is Face Validity? *Proceedings of the Second Annual Meeting of the South African Association for Research in Mathematics and Science Education. 22-26 January 2002. University of Natal, Durban KwaZulu-Natal, (III 364-370).*

Savage, M. 1998. Curriculum innovations and their impact on the teaching of science and technology. In Naidoo, P. & Savage, M. (Eds.) *African Science and Technology Education into the New Millennium: Practice, Policy and Priorities, (35-59).*

Savenye, W. C. & Robinson, R. S. 2001. Qualitative Research Issues and Methods: An Introduction for Educational Technologists. In Jonassen, D. H. (Ed.). 2001. *Handbook of Research for Educational Communications and Technology.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. (1171-1195).

Savery, J. R. & Duffy, T. M. 1995. Problem Based Learning: An Instructional Model and Its Constructivist Framework. *Educational Technology/September-October. (31-38)*

Scheffler, I. 1967. *Science and Subjectivity. Indianapolis.* The Bobbs-Merrill Company Inc.

Schlick, M. 1925. Philosophy of Organic Life. In Feigl, H. & Brodbeck, M. (Eds.). 1953. *Readings in the Philosophy of Science.* New York: Appleton-Century-Crofts, Inc.

Scott, P., Dyson, T., & Gater, S. 1987. *Children's Learning in Science Project.* A constructivist view of learning and teaching in science. Leeds: Centre for Studies in Science and Mathematics Education.

Senior, S. 1990. *Using IT across the national curriculum. Data handling key stage 2.* Kent: Owlet Books.

Shakeshaft, C. 1999. Measurement Issues with Instructional and Home Learning Technologies. [Online] Available: http://www.ed.gov/Technology/TechConf/1999/whitepapers/paper9.html [22nd July 2000].

Shepere, D. 1984. *Reason and the Search for Knowledge. Investigations in the Philosophy of Science.* Boston. D. Reidel Publishing Company.

Shymansky, J. A., Yore, L. D., Treagust, D. F., Thiele, R. B., Harrison, A., Waldrip, B. G., Stocklmayer, S. M., & Venville, G. 1997. Examining the construction process: A study of changes in level 10 students' understanding of classical mechanics. *Journal of Research in Science Teaching, 34,* (571-593).

Smith, M. G. & Simmons, P. E. (Eds.) 1992. Implementing a conceptual basis for teaching and learning genetics. Teaching genetics, recommendations and research. *Proceedings of a National Conference.*March 18-21, 69-78. Cambridge, Massachusetts.

Soriano, F. I. 1995. *Conducting needs assessment. A multidisciplinary approach.* Thousand Oaks: Sage Publications.

Spector, M. 1965. Models and Theories. In Brady, B. A. & Grandy, R. E. (Eds.) 1971. *Readings in the Philosophy of Science. Second Edition.* New jersey. Prentice Hall.

Splitter, L. J. 1991. Critical Thinking: What, Why, When and How. *Journal of Educational Philosophy and theory 23(1) (89-108).*

Sprinthall, N. A. & Sprinthall, R. C. 1990. *Educational Psychology. A Developmental Approach. Fifth Edition.* McGraw-Hill, Singapore.

Stake, R. E. 1994. Case Studies. In Denzin, N. K. and Lincoln, Y. S. (Eds.) (1994). *Handbook of Qualitative Research.* London: SAGE Publications. (236 – 248)

Stevenson, R. B. 1995. Action Research and Supportive School Contexts. Exploring the possibilities for Transformation. In Noffke, S. E. & Stevenson, R. B. (Eds.) 1995. *Educational Action Research. Becoming Practically Critical.* New York: Teachers College Press. (197-209).

Stockman, N. 1983. *Antipositivist Theories of the Sciences.* Boston. D. Reidel Publishing Company.

Stratford, S. J. 1997. A Review of Computer-Based Model Research in Precollege Science Classrooms. In Krajcik J. S. 1997. *Journal of Computers in Mathematics and Science Technology 1997 16(1), (3-23).*

Sugar, S. & Sugar, K. K. 2002. *Primary Games. Experiential learning Activities for Teaching Children K-8.* San Francisco: Jossey-Bass.

Sutton, C. 1996. Beliefs about science and beliefs about language. *International Journal of Science Education Vol. 18, No. 1, (1-18).*

Taiwo, A. A. 1995. *Fundamentals of Classroom Testing.* New Delhi: Vikas Publishing House PVT LTD.

Tamir, P. 1996. Science Assessment. In Birenbaum, M. and Dochy, F. J. R. C. 1996 (Eds.). *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge.* London: Kluwer Academic Publishers.

***The American Heritage Dictionary of the English Language, Fourth Edition,* 2000.** Houghton Mifflin Company.

The Department of Computer-Based Education, University of Cape Town. 2000. [Online] Available: http://www.uct.ac.za/projects/cbe/mcqman/mcqchp2.html
[21st November 2001].

The Education Policy Unit, University of Western Cape, 2000. *National Survey on Computers in Education in South African Schools. Bellville 7535.*

The Joint Committee on Standards for Educational Evaluation. 1994.
The Program Evaluation Standards, 2nd Edition.

The Maricopa Center for Learning & Instruction, Maricopa Community Colleges, 2000. What is assessment of learning? [Online] Available: http://www.mcli.dist.maricopa.edu/ae/al_what.html [20th March 2001].

Thelwall, M. 1998. Computer-Based Assessment: a Versatile Educational Tool. *Computers & Education 34 (2000)* (37-49).

Thompson, A. D., Simonson, M. R. & Hargrave, C. P. 1992. *Educational Technology. A Review of the Research. Revised Edition.* Washington Dc: Association for Educational Communications and Technology.

Thomson, N. & Stewart, J. 2003. Genetics Inquiry: Strategies and Knowledge Geneticists Use in Solving Transmission Genetics Problems. *Science Education 87, (161-180).*

Tinker, R. F. & Papert, S. 1988. Tools for Science Education. In Ellis, J. D. (Ed.) 1989. *Information Technology and Science Education, (1-23).*

Tobin, K., Tippins, D. J., & Gallard, A. J. 1994. Research on Instructional Strategies for Teaching Science. In Gabel, D. L. (Ed.). 1994. Handbook of Research on Science Teaching and Learning. New York. Macmillan Publishing Company.

Trochim, W. M. K. 2002. Research Methods Knowledge Base. [Online Textbook] Available: http://trochim.human.cornell.edu/kb/ [17th February 2003].

Tsai, C-C. 2000. Relationships Between Student Scientific Epistemological Beliefs and Perceptions of Constructivist learning Environments. *Educational Research Vol. 42 No. 2 Summer, (193-205).*

Turoff, M. 1995. Designing a Virtual Classroom. Presented at the International Conference on Computer Assisted Instruction ICCAI'95. March 7-10, 1995. [Online] Available: http://www.njit.edu/njIT/Department/CCCC/VC/Papers/Design.html [16th May 2001].

Twomey, E. and Miller, P. July 1996. *Computer-Based Assessment: An Introduction.* Life Science Educational Computing Vol. 7 [Online] Available: http://www.liv.ac.uk/ctibiol/lsec/July96.pdf [13th February 1999].

Vygotsky, L. S. 1962. *Thought and language.* Edited and translated by Eugenia Hanfmann and Gertrude Vakar. Cambridge, MA, NY: MIT Press; John Wiley.

Vygotsky, L. S. 1978. *Mind in society.* Cambridge, MA: MIT Press.

Wade, N. 2003. Watson and Crick, Both Aligned and Apart, Reinvented Biology. [Online] Available: http://www.nytimes.com/2003/02/25/science/25FATH.html?ex=1047312564&ei=1&en=791776f7c88f0880 [27th February, 2003].

Wartofsky, M. W. 1968. Conceptual Foundations of Scientific Thought. An Introduction to the Philosophy of Science. London. Collier-Macmillan Limited.

Watson, B. 2001. Key Factors Affecting Conceptual Gains from CAL Materials. *British Journal of Educational Technology. Vol. 32 No. 5, (587-593).*

Weiss, C. H. 1998. Evaluation. *Methods for Studying programs and Policies.* New Jersey: Prentice Hall.

White, J. A. & Purdom, D. M. 1996. Viewing Modern Instructional Technology Through Conceptions of Curriculum. *Educational Technology Review. International Forum on Educational Technology Issues and Application. Autumn 1996, No. 6, (5-9).*

White, R. T. & Tisher, R. P. 1986. Research on Natural Sciences.
In Wittrock, M. C. (Ed.). *Handbook of Research on Teaching. Third Edition.* New York: Macmillan Publishing Company.

Willis, J. 2000. The Maturing of Constructivist Instructional Design: Some Basic Principles That Can Guide Practice. *Educational Technology/January-February, (5-16).*

Wilson, B. G. 1995a. "Metaphors for instruction: Why we talk about learning environments." *Educational Technology, 35(5), (25-30).* [Online] Available: (http://carbon.cudenver.edu/~bwilson/wils95) [20th July 2002].

Wilson, B. G. 1995b. Maintaining the Ties Between Learning Theory and Instructional Design. *Paper presented at the meeting of the American Educational Research Association, San Francisco, March 1995.* [Online] Available: http://www.cudenver.edu/-bwilson [8th October 2002].

Wilson, B. G. 1996. *Constructivist learning environments: Case studies in instructional design.* Englewood Cliffs NJ: Educational Technology Publications.

Wilson, B. G. & Cole, P. 1991. Cognitive Teaching Models. A Review of Cognitive Teaching Models. *Educational Technology Research and Development, 39(4) (47-64).* Also [Online] Available: http://www.cudenver.edu/~bwilson/cogapp.html
[30th January 2002].

Wilson, B. G. & Cole, P. 1996. Cognitive Teaching Models. [Online] Available: http://www.cudenver.edu/~bwilson/hndbkch.html [17th July 2001].

Wilson, B. G., Jonassen, D. H., & Cole, P. 1993. Cognitive approaches to instructional design. In Piskurich G. M. (Ed.), *The ASTD handbook of instructional technology,* (21.1-21.22). New York: McGraw-Hill. Also [Online] Available: http://www.cudenver.edu/~bwilson [16th October 2001].

Winn, W. 1993. A Conceptual Basis for Educational Applications of Virtual Reality. [Online] Available: http://www.hitl.washington.edu/projects/education/winn/winn-R-93-9.txt#imm [15th October 2001].

Winn, W. 1997. The Impact of Three Dimensional Immersive Virtual Environments on Modern Pedagogy. [Online] Available: http://www.hitl.washington.edu/publications/r-97-15/ [16th November 2000].

Wise, S. L. & Plake, B. S. 1990. Computer-Based Testing in Higher Education. *Measurement and Evaluation in Counselling and Development / April 1990 / Vol 23.*

Witkin, B. R. & Altschuld, J. W. 1995. Planning and conducting needs. A practical guide. Thousand Oaks: Sage Publications.

Wolcott, H. F. 1992. Posturing in Qualitative Research. In Le Compte, M. D., Millroy, W. L., and Preissle, J. (Eds.). *The Handbook of Qualitative Research in Education.* New York: Academic Press Inc.

Wollman, W.T. 1990. Applying "Cognitive Conflict" Strategy for Conceptual Change – Some Implications, Difficulties, and Problems. *Science Education 74(5), (555-569).*

Wood-Robinson, C., Lewis, J., Leach, J., & Driver, R. 1997. Scientific Literacy and the School Curriculum: Rationale, Design and Methodology for an Investigation of Young People's Understandings of Genetics and Their opinions on, and Attitudes to, New Gene Technologies. *Meeting of the Association for Research in Mathematics, Science and Technology Education. 22-26 January 1997. University of Witwatersrand, Johannesburg, (181-187).*

Yeaman, A. R. J., Hlynka, D., Anderson, J. H., Damarin, S. K. & Muffoletto, R. 2001. Postmodern and Poststructural Theory. *Handbook of Research for Educational Communications and Technology.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. (253-295).

Yore, L. D. June 2001. What is Meant by Constructivist Science Teaching and Will the Science Education Community Stay the Course for Meaningful Reform? Electronic Journal of Science Education Vol. 5 No. 4 - June 2001. [Online] Available: http://unr.edu/homepage/crowther/ejse/yore.html [18th March 2002].

Young, J.D. 1996. The effect of Self-Regulated Learning strategies on performance in Learner Controlled Computer-Based Instruction. *Educational Technology, Research and Development. 44 (2), (17-27).*

Appendix I: Tests

A. Test 1 set by the science teachers at East London

Table of specifications (adopted from Taiwo, 1995: 41)

Cognitive Domain	No. of items		Science topics						
			Roots	Cell structure	Acids	Electricity	Electrostatics	Decomposition	Force / Pressure
Knowledge	7	14	(9), (10)	(8)	(6)	(2)		(5), (7)	
Comprehension	1	2					(1)		
Application	-	0							
Analysis	2	4							(3), (4)
Synthesis	-	0							
Evaluation	-	0							
Total	10	20	2	1	1	1	1	2	2

1. *The following statement on electrification is not true:*

a. It occurs according to the Law of Conservation of charge
a. It occurs through friction
a. It occurs when electrons are transferred from one object to another
a. It is the result of creation of some charges and the destruction of others

2. The factor that does not influence the resistance of a conductor is _____

a. Mass b. Temperature c. Diameter d. Type of material

3. The weight of a 12 Kg. Mass on the earth is more or less __________

a. 120 Newton b. 12 Newton c. 120 Joules d. 1.2 Newton

4. A man stands on a bathroom scale. He now lifts one foot. Which one of the following statements is true?

a. The force exerted by the man on the scale will double
b. The scale will then register a greater reading
c. The pressure exerted by the man on the scale will double
d. The man's weight will decrease

5. The correct formula for Ammonia is __________

a. N_3H b. NH_4 c. NH d. NH_3

6. The common name of Calcium Hydroxide is __________

a. Limestoneb. Slaked lime c. Unslaked lime d. Quick lime

7. The equation $CuCO_3 \rightarrow CuO + CO_2$ represents

a. The decomposition of Copper Carbonate through heating;

b. The combustion of Copper Carbonate

c. The neutralisation of Copper Carbonate

d. The reaction of a carbonate with an acid

8. The part of the plant cell filled with cell sap is called __________

a. Cytoplasm b. Vacuole c. Nucleus d. Cell wall

9. The region of the young root of a plant where cell division takes place is called:

a. Meristematic region

b. Root cap

c. Region of elongation

d. Mature region

10. __________ has adventitious roots.

a. Geranium

b. The bean

c. Wheat

d. The plant of carrot

Memorandum

1) d 2) a 3) a 4) c 5) d

6) b 7) a 8) b 9) a 10) c

B. *Diagnostic Test 2: Derived from Test 1, on electricity (left on the school LAN)*

---

1. If one increases resistance, current will

a. Decrease b. Remain the same
c. Increase d. It will not affect current (2)

2. The unit of electric charge is----

a. Coulomb b. Ohm
c. Ampere d. Volt **(2)**

3. The Ohm is the unit of -----

a. Resistance b. Charge
c. Current d. Potential difference **(1)**

4. If current increases, resistance will-----

a. Remain the same b. Increase c. Decrease (2)

5. The effective resistance of two resistors of 2 Ohms each in parallel is___(1)

6. Battery A is marked 8 Volts and sends out 2 Coulombs in one second, Battery B is marked 4 Volts and sends out 4 Coulombs in one second, and Battery C is marked 16 Volts and sends out 2 Coulombs in two seconds. From the list below, select FOUR statements that are true.

a. Coulomb in Battery C has more energy than a coulomb from Battery A and Battery B. **(3)**
b. Battery C is pushing out the lowest current (1)
c. Battery B could be experiencing the lowest resistance (2)
d. Battery B is the smallest in size because it is producing the lowest Volts **(0)**
e. Battery B will be most powerful after an hour **(-1)**
f. All batteries produce the same amount energy in one second. **(0)**
g. The three batteries are producing the same power **(4)**
h. The current is highest through Battery C because it has the highest Volts **(0)**

7. In a higher resistance fewer charge pass through, and so the potential difference across that resistor will have to be low.

a.	Strongly agree	(-2)
b.	Agree	(-1)
c.	Disagree	(1)
d.	Strongly disagree	(2)
e.	Neither agree or disagree	(0)

8. If a current of 2 Amperes flows through a battery with a potential difference of 4 Volts, in 3 Seconds. The amount of charge passing through the battery is--- (2)

9. Potential difference is about energy, which one Coulomb gains or loses between two points in an electric circuit. It has nothing to do with the resistance in the circuit.

a. Strongly agree (-2) b. Agree (1) c. It is partly true and partly wrong (2)
d. Disagree (1) e. Strongly disagree (-2)

10. Siphokazi connects a resistor of 12 Ohms in parallel with another resistor of 6 Ohms. Akhona connects a resistor of 12 Ohms in series with a resistor of 6 Ohms. Olwethu has a resistor of 4 Ohms, but decided to heat it up slightly over a flame. Select THREE possibilities.

a.	Siphokazi has made the lowest resistance	(3)
b.	Olwethu's resistance could be higher than that of Siphokazi	(3)
c.	Charge will be faster in Siphokazi's circuit	(2)
d.	Olwethu's resistance is equal to that of Siphokazi	(0)
e.	Akhona's resistance is less than that of Siphokazi	(-1)
f.	The strength of the resistors will depend upon the current through them.	(-2)

11. A volt can be described as -----

a. A Joule per Coulomb
b. A Joule per second
c. A Coulomb per second
d. A Watt per second (1)

12. From the definition of potential difference, it means that if 2 Coulombs of charge flow through a battery of 1.5 Volts, the coulombs will -----. Pick the TWO correct answers.

a.	Gain 3 Joules altogether	(2)
b.	Gain 1.5 Joules each	(1)
c.	Gain 3 Watts each	(0)
d.	Gain 3 Newtons altogether	**(0)**

Appendix II: Technical validation of the CAT by an ID expert

Question: How technically sound is the test?

Responses from the Instructional Designer Expert (Technical checklist)

Date of review : 11th November 2001

Name of programme : CAA (Question Mark) for schools

Description of programme : Physical Science CAT for Grade 10

Loading the CAT

1. Any problems encountered when
   a. Loading the test? If yes explain briefly: *No*
   b. Accessing the different questions once loaded? If yes, explain briefly: *No*

Using the CAT

2. Can the user stop the CAT by mistake? *No*
3. Are there obvious faults in the running of the CAT? *No*
4. If the CAT fails to run properly, can the error be easily traced? *Yes*
5. Does system alert the user to corruption of files or problems? *N/A*
6. Are dialog boxes for error messages provided? *No*
7. Is the desktop visible when the test is running? What proportion of the screen is used? Does it allow re-sizing? *No. Full screen is used.*

Design considerations

8. Ideally the learners should be able to browse through the test easily – is this possible? *Yes*
9. Is it easy to get lost in the test? Explain. *No*
10. Is the status shown all the time and in the same position? e. g. where to go next, options available etc *Yes*
11. Can the leaner/teacher easily recover data lost? *Don't know*
12. How many colours are used? *Too many*
13. Does the user consider the colours friendly? (Any cultural or social associations with the colours used?). *None for European culture but colour schemes should be more standard.*

14. Is typing (if any) part of the educational process? Are users required to be proficient typists? *No*

15. Where is the focus of the test – keyboard, screen etc? *Uses both*

16. How many keys are required to control the test (Or can the test be run by the use of the mouse only?

A mouse, except two questions, which require typing in, can run it.

Presentation:

Key; SA = Strongly Agree; A = Agree; U = Undecided; D = Disagree
SD = Strongly Disagree; NA = Not Applicable

17. The overall presentation is attractive

SA	**A**	**U**	**[✓ D]**	**SD**	**NA**

18. The use of colour is suitable for the subject

SA	**A**	**U**	**[✓ D]**	**SD**	**NA**

19. The use of language is suitable for my learners (Its terminology and notation are not confusing)

SA	**[✓ A]**	U	D	**SD**	**NA**

Operation

20. The CAT is easy to operate

[✓ SA]	**A**	**U**	**D**	**SD**	**NA**

21. A child with little experience of computers can use the test

[✓ SA]	**A**	**U**	**D**	**SD**	**NA**

22. It is easy to modify the CAT if the need arises

SA	**A**	**[✓ U]**	**D**	**SD**	**NA**

Printing:

23. Does the test allow printed output? *No*

24. Does the printer work properly with the test? *N/A*

25. Are there cheaper alternatives? *Possibly*

Further comments made by the expert

i. *There are* ***several errors*** *of science in the test that should be corrected before the test is used.*

ii. *Physical Science test needs a more standard interface as the* ***OK button moves position erratically*** *on the screen.*

iii. ***Colour should be selected for function rather than decoration****. E.g: one colour scheme for electricity.*

iv. *Questions would better be* ***moved away from the extreme top of the screen****.*

v. *Font used was pleasant and friendly but the* ***same font should be used throughout the question****.*

vi. *Questions could benefit from* ***more graphical representation.***

vii. *Use of the system is intuitive but should be taken with physical installation, as errors on disks can be disastrous to a lesson plan.*

Appendix III: Responses from the interviews about CAA applied to two teachers

---

T1 Science Teacher

T2 Science Teacher

Note: Exponent numbers are identities of themes

Subsidiary questions

1. From your observations of your learners doing the CAT and the data that can be obtained, in what way would CAA help you with your teaching?

T1 *It is **interesting** ^{1} to the students. Students are **excited** ^{1} about the CAT. It opens their mind to do it. The CAT is **helpful** ^{2} in that*

- *I don't have to give out sheets of paper for the test*
- *It saves me marking time, and offers immediate feedback*
- ***Tells me the errors and mistakes that learners make so that I can concentrate on that chapter*** ^{7}

T2 *– The test can **be done unsupervised** ^{3}, especially **when you have a few teachers** ^{4}. It is similar to Master maths, except that Master Maths also teaches. I **can cover more ground in big classes** ^{4}. For example, I **could attend to only those students who don't understand** ^{5}. It also encourages self-learning.*

2. How would you fit CAA in your assessment?

T1 *– After every chapter; **for revision** ^{6} to see how they understand what I have taught.*

T2 *– Can be given at end of a chapter **for revision** ^{6}, and as a class test or a combination of the two.*

3. What could make CAA educationally attractive to your learners?

T1 *– They are **more relaxed** ^{3}*

T2 *– Learners didn't say much. They found it **convenient** ^{2}, for example, it shows the time left and therefore indicates how fast they need to work.*

4. What can discourage you from using CAA?

T1* - *Nothing

T2 *– Lack of another laboratory and computers:* ***computers are not enough*** ^{7} *because those available are used for so many things.*

5. Have you observed any problems with your learners using the CAT?

T1 – No T2 – No

6. Would you advise your school to use CAA?

T1 *– Yes, definitely*

T2 *– No. It* ***requires a lot of money*** ^{11}*. It may be difficult to convince the school to spend over R 250 000 necessary for another computer laboratory and computers.*

7. Are these options clear to the user at all stages?

T1 *– It is* ***clear*** *T2 – Very nice.*

8. The main objective of this particular research was to find out the diagnostic value of CAA. Can you identify learners' problems from the data provided by CAA?

T1 *– Yes.*

T2 *– It depends on the teacher and the questions. For example, it depends on what you want to achieve so that you* ***structure the questions*** ^{10} *for that. It would also* ***require more questions*** ^{12} *to make sure about the problems learners have.*

9. How well can CAA be used to provide teachers with diagnostic information on learners' current understanding of the science concepts at Grade 10 level?

T1 *– It* ***can show where the problem*** ^{5} *is,* ***what kind of mistakes*** ^{7}*,* ***how the learner thinks*** ^{8}*, and why something is not clear.*

T2 *– It is okay*

11. In your opinion, what are the major strengths of CAA as used in this case?

T1 *– It* ***encourages the learner to do the test*** ^{2} *and to find out more by him/herself*

T2 *– It* ***takes the burden off the teacher*** ^{2}*, and so* ***saves time*** ^{2} *for the teacher to* ***help the more needy students*** ^{5}*.*

12. In your opinion, what are the major weaknesses of CAA as used in this case?

T1 *- None*

T2 *– It* ***requires finance*** [11]*, infrastructure, and electricity. It also requires the teachers to have* ***knowledge of how to set questions on the software and then administer the questions to their students*** [10]*. I think it also matters whether the teacher knows* ***how to interpret and use the data*** [9]*.*

13. What changes would you recommend on this product?

T1 – NoneT2 - None

Appendix IV: Learners' responses to the fairness and value of CAA

(Learners were coded; for example, Learner 1 is L1, learner 2 is L2, etc.)

Note: Exponent numbers are identities of themes

Name and answer to question 2 above	**Time to complete / Min**	**Mark s/ %**
L1 - *It was **quite good** ^{1}. Some of them were very **difficult** ^{4}*	18:33	33 %
L2 - *The test to me was **interesting** ^{2} in a way. In the way that it **made me think a lot** ^{3} and **gave me a chance to do what I feel**. It made me realise how much I really know about science but the **answers were not revealed at the end. I wanted to see where I went wrong**.*	14:16	33 %
L3 - *It was a bit **difficult** ^{4} because there were things I did not know. And I liked it. It was **interesting** ^{2}. But it was great!!!*	19:55	37 %
L4 - *I found the test **interesting** ^{2}. It was very challenging. **Needed you to know your facts**.*	19:37	37 %
L5 - *It was challenging. It was not that bad. The questions were realistic. It helped me **revise** ^{5} because I have done all those kind of questions. The questions only want you to use your common sense. I **enjoyed** ^{6} it very much.*	12:41	40 %
L6 - *Well, actually the test was fine and bit **difficult** ^{4}. It's good because you're testing your knowledge, and **also learn from your mistakes by reckoning and giving right answers**.*	10:48	42 %
L7 - *It was **not simple** ^{4} and I don't think I've got the highest mark as I thought I was. Some of the questions were simple. It was good to us. I know that some of us know something out of it. I hope we will do most of these tests in other subjects*	11:11	44 %

***L8** - It was **very nice**[1] and I enjoyed it very much but it was **difficult**[4] in some questions and **I wish I could write it again**.*	19:09	50 %
***L9** - This test shows how you are experienced and it shows how much did you **revise**[5] on your science and biology. When I was writing this test I was scared because I don't know what I was going to do. It was **very easy**[7] if you study and **it was the bright ways to be examined**.*	11:06	50 %
***L10** - First of all I can say this test gives me a clue for my physical science exams. It **encourages me to study**[8] more and more. It tells me that there is nothing difficult and there is nothing easy only you that you can make things difficult or easy for you.*	15:46	50 %
***L11** - Although it took us by surprise, we did not expect this. I thought I was gonna experience something and learn something. **It was challenging**[4] and very **stressing because I was not prepared**[9]. I do know these things but man I was surprised. Everything seemed to have flushed off my mind. Sorry Mrs Spates if I disappointed you. Don't take me wrong I did learn.*	14:41	62 %
***L12** - I think it **is ok**[10]. It should be improved on Biology side for example, add more questions. I think there should be Maths questions added to the test. Basically I think it is great.*	16:36	62 %
L13 *No response*	10:03	73 %
***L14** - The test was of the things we know but it was put in a **difficult**[4] **manner so that we can not exactly understand the questions**. Although it was difficult, we could answer the questions, and I feel it is a good way of testing students*	13:23	75 %

Appendix V: A sample of learners' responses concerning Zadarh

2. What events do you remember when using this material?

School 2 - girls

S2-g1	I remember opening the cupboard in the store room where we had put back the cylinder
S2-g2	I remember where we had shapes and unable to complete them. When we could not get out of the room with a piano.
S2-g3	There was fire and we needed to put it out. We played the keyboard using instructions given in the notes.
S2-g4	The playing of the keyboard; lift; structures; tokens for photosynthesis and for respiration
S2-g5	The store room on fir. Getting Oxygen. Interpreting the keyboard notes, matching the enzymes. Having to find those two coins with the code.

School 3 - girls

S3-g1	Getting carbon dioxide to put out the fire in the store room. Running around in circles (getting lost).
S3-g2	When we had to open the safe by calculating the molecular mass. I also remember when we struggled to find the door to turn of the fire.
S3-g3	The time which we have to look for the keys that were in pot plants. We find the oxygen by first adding carbon dioxide with animal cell then to the pressure.
S4-g4	Getting the gas tank filled with oxygen and distinguishing of the fire in the same room.
S5-g5	Trying to refill the oxygen tank. Finding one of the doors.

School 2 - boys

S2-b1	The lock and key enzyme action
S2-b2	The gas cylinder were a bit tricky because we kind of forgotten the light
S2-b3	Fire extinguishing and gas tank re-loading
S2-b4	The air cylinder that required air and the photosynthesis experiment
S2-b5	The place where you find glucose and water, and then you have to calculate the code that you should use as a password

School 3 - boys

S3-b1 We solved the safe combination of the weight of carbon, Hydrogen, and Oxygen

S3-b2 The collection of carbon dioxide from carbon dioxide from oxygen and cellular respiration. The deciphering of codes.

S3-b3 Basics of respiration and photosynthesis. Putting fire out with carbon dioxide. Breathing in oxygen.

S3-b4 The filling of the tank with carbon dioxide to put out the fire. Trying to figure out the coding to open the entrance.

S3-b5 We could not solve the piano problem. We spent more than 15 minutes trying to work it out. Wew solved and managed to put out the fire.

3. What information do you remember from using this material?

School 2 - girls

S2-g1 The process of respiration and photosynthesis

S2-g2 That whatever is there goes hand in hand. I.e., if you had a key, you probably go for the door.

S2-g3 What is needed for cellular respiration

S2-g4 Cellular respiration - Carbon dioxide is released. Photosynthesis is opposite to respiration.

S2-g5 The process of photosynthesis and process of respiration. I must get the Oxygen so as to get in the store room. You must be smart like have your keys and know where you'll use them.

School 3 - girls

S3-g1 That you need carbon dioxide to put out fire. Glucose produces pyruvic acid

S3-g2 2 glucose produce 4 pyruvic.

S3-g3 The fact that you have to calculate add all the molecular weight.

S3-g4 Oxygen and light is needed for photosynthesis where as light is not needed.

S3-g5 I learnt more about photosynthesis. The molecular weight.

School 2 - boys

S2-b1	That only the exact part of the puzzle will fit the enzyme action
S2-b2	We gained more information in respiration and photosynthesis and learned that Oxygen is used in respiration
S2-b3	Photosynthesis and enzymes
S2-b4	The molecular form of compounds, and that respiration requires oxygen
S2-b5	That under photosynthesis for plants to photosynthesise they need light, water, and carbon dioxide. Light from a light bulb do not fit in photosynthesis

School 3 - boys

S3-b1	Through cellular respiration you can convert carbon dioxide to oxygen or vice versa.
S3-b2	Plants make oxygen from carbon dioxide and animals make carbon dioxide from oxygen. The formula of glucose.
S3-b3	Photosynthesis occurs during the day. Humans take in oxygen and take out carbon dioxide
S3-b4	Photosynthesis requires light and carbon dioxide. Chemical formula of glucose.
S3-b5	One can only move to an object by moving horizontally or vertically

4. Which part in the process of using the programme teaches you most?

School 2 - girls

S2-g1	The correction of puzzles (the enzymes) – its mostly about life itself
S2-g2	Enzymes where we had to put shapes. The enzyme par and finding the code
S2-g3	Playing the piano
S2-g4	The working of enzymes. The difference between the process involved.
S2-g5	The clear structures of chloroplasts and mitochondrion

School 3 - girls

S3-g1	Calculating molecular masses
S3-g2	When you try solving the problems of the game.
S3-g3	When trying to solve the problem.
S3-g4	The molar masses and calculations.
S3-g5	I learnt more when trying to solve a problem especially the piano notes.

School 2 - boys

S2-b1 I think the best was when we tried to gain oxygen from the process of photosynthesis

S2-b2 Teaches me more in respiration and photosynthesis.

S2-b3 The puzzle-solving

S2-b4 The chapter of photosynthesis and respiration is well demonstrated than in books

S2-b4 Finding the combination of the safe – you have to work hard

S2-b5 Using tokens to open something

School 3 - boys

S3-b1 Through interacting and by seeing how things we learn about work actually teaches us a lot

S3-b2 The deciphering of codes teaches us a lot.

S3-b3 Mitochondria supply energy needed for respiration. Chloroplasts supply energy needed for photosynthesis.

S3-b4 The manner in which the information is set out. I.e., fun, exciting

S3-b5 The interaction between the user and the program. The amount of knowledge or understanding at your biology.

5. How does this programme teach you?

School 2 - girls

S2-g1 From the enzymes part you learn how to deal with different situation in life.

S2-g2 It teaches you to link the pictures and the information from class. You need to know exactly where you are heading.

S2-g3 It needs 100% focus and logic

S2-g4 It needs a person to think very hard to figure out new to handle a situation at the same time it teaches you biology. To be attentive and think smartly to keep going

S2-g5 It tells/shows you how much you know, how much you understand. Shows you your flexibility.

School 3 - girls

S3-g1 How to basically solve problems with not much information or instructions.

S3-g2 It teaches me more on how to plan and also exercise my memory.

S3-g3 By the fact that it is tricky. You have to search a lot.

S3-g4 It challenges the mind. It requires you to research and ask why and how.

S3-g5 You have to be patient and a great deal of concentration is of essence. It teaches you to actually use your mind and to notice more of your surroundings.

School 2 - boys

S2-b1 It is an excellent teacher, it is fun and it is very organised

S2-b2 You discover a lot. It teaches me not panic and to also double-check

S2-b3 You discover a lot of things while playing

S2-b4 That its easy to learn by pictures than reading a textbook and to be very observant inside the rooms

S2-b5 To remember things

School 3 - boys

S3-b1 It teaches us letting us search and find information for ourselves

S3-b2 It incorporates fun with learning, it isn't a bore or tedious.

S3-b3 It teaches by concentrating on basics of respiration and photosynthesis. It also teaches through puzzles.

S3-b4 Introduces you to computer graphics and biology

S3-b5 *No answer*

6. Can you advise what you would add or take away from this programme?

School 2 - girls

S2-g1 I think it needs change in the direction where you have to go on angle. May be it is better when you want to go to an object go straight to it and when you want to achieve an obstacle, then you can go on angles.

S2-g2 I think more information in the dark room so you have an objective whether you die or not (we want to get anything that would encourage us to score even more).

S2-g3 Moving towards an object

S2-g4 The movements need to be straight forward (not at 90^0). The thinking must be done when solving problems not when.

S2-g5 The place must look more like a lab. So that you can feel the spirit of science, chemicals, etc., and anything concerning biology.

School 3 - girls

S3-g1 I think it should be done in levels and get a particular score to get to the next level. Every time you get or solve a problem, 'one' should get an instruction on what to do next. I think it is too organise and too formal, if may be it was a bit informal then may be it would be more fun.

S3-g2 I'd prefer it or love it more if the whole view of the room was shown at the same time.

S3-g3 Its perfect for me. I personally do not suggest any changes in the game.

S3-g4 There isn't much of any change except choice of words like I'd prefer compass instead of direction.

S3-g5 They have to show more of the rooms.

School 2 - boys

S2-b1 Add some more doors to increase the confusion

S2-b2 Shape of the rooms

S3-b3 The rooms – they are the same, so I would change rooms.

S3-b4 Room shapes and the 90^0 direction change to at least 30^0

S3-b5 The shape of rooms is really confusing. The furniture inside the rooms should be different because they look akin in every room. Move in all directions and use buttons to move

School 3 - boys

S3-b1 The camera angles and information layout

S3-b2 More clues be given to the player

S3-b3 Camera angles. To be more informative than fun.

S3-b4 Camera angles. Information.

S3-b5 *No answer*

NATAL UNIVERSITY

A manual for

EVALUATING EDUCATIONAL COMPUTER PROGRAMMES

Johnnie Wycliffe Frank Muwanga-Zake

2004

FOREWORD

A word to designers and evaluators

The following require the attention of designers:

- Designers should provide a guide or manual on how a programme should be used
- The evaluator must understand clearly the aims of a programme, and the curriculum in which it is supposed to be used
- Teachers and learners should participate in the evaluation since educational programmes aim at solving problems in teaching or learning
- For computer educational programmes, it is preferable that:
  - c. The claims made by the designers with regard to how the programme assists learning are made explicit
  - d. The evaluator has practiced teaching and has a qualification in education

Notes on the evaluation

- The utility should be accompanied by a user's manual. This could be within the programme.
- Teachers and/or learners fill in some parts of the questionnaire and then are interviewed for some other parts of the questionnaire, after using the programme.
- It is important that principal, teachers, and learners consent to the evaluation. The DoE should be consulted if the evaluation is conducted in a government school
- The observation schedule as well as interviews could be supported by pre-test/post-tests. However, it should be noted that some qualitative aspects of utilities are not easily captured by traditional class tests.
- The observation schedule attempts to depict the scene and context of the application of the programme. With consent, a video record of the proceedings is recommended.
- It might be essential that if teachers are used as evaluators, teachers and/or their institutions benefit from the exercise. This may be in form of training on evaluation, giving them a utility free of charge, or remuneration.

Using the instrument

You are advised to modify the evaluation instrument to suit your needs. For example, you may find that some of the questions do not apply to your evaluation.

Your participation

The developers of this instrument welcome submission on its weaknesses, and in fact would be pleased to receive your criticism for further development of this instrument.

NOTE: PLACE A TICK (√) IN THE APPROPRIATE BOX WHERE READY ANSWERS ARE PROVIDED

Please contact me if you get problems or comments with this evaluation instrument

J. W. F. Muwanga-Zake:
Office - 031 2603418
Cellular - 0837521534
e-mail jmuwanga@lycos.com

PART A

EVALUATOR'S EXPERIENCES & QUALIFICATIONS

---

1. YOUR QUALIFICATIONS AND EXPERIENCE IN TEACHING THE SUBJECT CONCERNED

Grades in which you teach		Teaching qualification	
Number of years teaching		Highest level to which you studied	

2. YOUR QUALIFICATIONS AND EXPERIENCE WITH COMPUTERS

2.1 Your qualifications in using computers

____

2.2 For how long have you used computers in education?.

Never	Less than a year	Up to 5 years	6-10 years	11-15 years	Over 16 years

3. YOUR PREVIOUS EXPERIENCES WITH EVALUATION (State what you have evaluated before)

____

____

____

PART B

USER-PROGRAMME INTERACTIONS

---

4. TECHNICAL ISSUES

Place a (√) in the box of your choice

Note: Please make sure that the hardware and software of the computer you are using meets the requirements of the programme. Some problems happen because the computer is not suitable for the programme.

4.1 Skills / knowledge required

What special technical skills and knowledge are necessary in order to use this programme effectively?

Skill / knowledge	**Yes**	**No**
E.g. Typing skills		

4.2 Compatibility of the programme with school computers

	Immediate: Simple set-up, access, or understanding of controls (less than 10 minutes).	Moderate: Some attention needed to set-up, access or understand controls (About 10 minutes).	Significant: Complex set-up, access, or understanding of controls (more than 10 minutes).
Installation on the computer			
Starting the programme on the computer, and time required to access software, navigate menus and begin using the programme.			
Pre-Lesson set up. The time taken to become familiar with the programme or to explain the programme to somebody.			
Response of the programme to inputs			

4. 3 Help and documentation

Is there guidance, help or supporting documentation? The following may help you consider the above question:

a. Are the loading and running instructions clear?

Yes		Some of it		No	

b. Is sufficient information given to enable the user to know what the software does, and how it behaves without having to run the software?

Yes		Some of it		No	

c. Is how you move around in the programme clearly explained or marked?

Yes		Some of it		No	

d. Help. Is there a help option in a manual or on-screen to explain technical points, menus and icons?

Yes		Sometimes		No	

4.4 Design and navigation

How easy is it to brose through the programme? The following may help you consider the above question.

a. Is the vocabulary in the menus understandable to intended users (E.g., learners)?

Yes		No	

b. Are the icons useful and can they be easily selected by a mouse click?

Yes		Some of it		No	

c. List the icons that are not responsive

____

d. Can parts of the programme sequence be by-passed if desired?

Yes		Some of it		No	

e. Can you get in to and out of parts of the programme easily?

Yes		Some of it		No	

f. Can you restart where you left off?

Yes		Some of it		No	

Please mention parts where you had problems in moving around.

9. Can you make notes using the computer (E.g., using 'Word') whilst using the programme?

Yes		Some of it		No	

h. Can you select and print text or diagrams you want?

Yes		Some of it		No	

i. Does the programme keep records of performance?

Yes		Some of it		No	

j. Can the teacher access learner performances and identify each learners progress?

Yes		Sometimes		No	

k. What complaints do you have about the colours and graphics?

l. What complaints do you have about sound effects and music?

4. 5 Level of use

a. Can the level of activity be set?

Yes		Some of it		No	

b. Can the programme be used without much help from the teacher? E.g.:

i. Does the software provide a tutorial?

Yes		Some of it		No	

ii. Can learners find specific information or activity easily (without assistance)?

Yes		Some of it		No	

4.6 What questions or suggestions do you have for the designer ?

PART C

CURRICULUM ISSUES

---

5. CURRICULUM ISSUES

The programme must provide skills and knowledge that are relevant to users (E.g., learners and/or teachers). This section establishes the relevance of the programme.

Requests

a. Take time to go through the programme carefully as much as possible
a. Give your genuine opinions without reservation – no one will be offended by your criticism

5.1 Overview of teaching with the programme

These could include subject related skills, critical thinking, problem solving skills, generating hypotheses and testing them, application of number, etc.

a. List the benefits (from among those above or similar) of using this programme that are not as easily achievable under normal teaching in a class

b. List the weaknesses (from among those above or similar) of the programme for classroom use

c. State the range of Grades at school which could use this programme (E.g., Grade 10-12)

d. Is the quality of the content acceptable? Consider the following aspects:

   i. Accuracy. List the errors you have encountered (state that they are too many if this is the case)

   ii. Spelling. List the spelling mistakes (state that they are too many if this is the case)

iii. *Is there violence, adult language, or themes that are inappropriate to the intended users? Please list them.*

iv. *Is there any evidence of bias? E. g., cultural, gender or racial bias. Indicate in which way.*

5.3 Content

a. If the programme deals with content, which of the topics or concepts the programme purports to teach are not well-covered?

b. For what knowledge /content would you specifically recommend this programme?

c. Is the extent of the content sufficient and appropriate for the target audience? Is information sufficiently detailed or is information too detailed? Please explain with one example.

d. Is the information arranged logically? E.g., topics follow each other as a succession of developing ideas, as opposed to randomly linked material. Suggest a better logical sequence.

6. TASKS AND EXERCISES (ONLY FOR PROGRAMMES IN WHICH TASKS ARE GIVEN)

6.1 It is desirable to incorporate in a programme tasks or problems to solve. Are the tasks useful or relevant to the user?

Yes		Some of it		No	

6.2 Can you complete the tasks by only using the information provided in the programme?

Yes		Some of it		No	

6.3 For which tasks does the user need to use reference outside the programme?

6.4 Where exercises are offered on screen; Are these exercises easily and reliably accessed (or 'hidden' in the programme)?

Yes		Some of it		No	

6.5 Is there a logical sequence of tasks (E.g., Do the tasks or exercises become progressively more difficult?)

Yes		Some of it		No	

6.6 Is feedback given to reinforce accurate answers, and to correct wrong answers?

Yes		Some of it		No	

7. INTERACTIVITY AND ENJOYMENT

7.1. Is the programme genuinely interactive? Can pupils create a combination of information, which was not there before? E.g., help them to introduce their own ideas?

Yes		Some of it		No	

Provide an explanation if necessary

7.2. For programmes with virtual environments

a. Where the programme simulates a real world environment, do the laws governing actions and consequences, and the behaviour of individual elements, follow accepted models or rules related to the same real world situation?

Yes		Some of it		No	

b. Do the skills practised in the virtual environment match those that would be required in the physical world?

Yes		Some of it		No	

7.3 Would users enjoy using the programme? Enjoyment implies that users become absorbed into the programme.

Yes		Some of it		No	

a. Which aspects have you enjoyed? (One example is enough)

b. Can you obtain such enjoyment by other means in class?

Yes		Some of it		No	

c. Which aspects of the programme are boring? I.e., parts where users lose interest in the programme.

8. SPECIAL NEEDS

8.1. Can text be spoken (for those with disabled sight)?

Yes		Some of it		No	

8.2. Can the size and colour of text be altered for the visually impaired?

Yes		Some of it		No	

8.3 Comment on other ways by which disabled people could use this programme

9. YOUR CONCLUSION

This step establishes whether a programme contributes towards solving problems at school. Would you use **THIS PROGRAMME?** Please list in simple statements what you consider to be the most serious problems in teaching/learning the subject, which this programme could solve

PART D

EVALUATION AGAINST HOW THE PROGRAMME REPRESENTS THE NATURE OF A SUBJECT AND WHETHER IT CAN ACHIEVE THE DESIRED OUTCOMES

(This is designed for science in South African schools – use outcomes of the subject the programme is about)

Does the programme support this? Place a tick (√) in the appropriate box

10. LEARNING OUTCOME 1: SCIENTIFIC INQUIRY AND PROBLEM-SOLVING SKILLS (SCIENCE PROCESS SKILLS):

Key Skills	The learner is able to use process skills, critical thinking, scientific reasoning and strategies to investigate and solve problems in a variety of scientific, technological, environmental and everyday contexts. *Learners' understanding of the world will be informed by the use of scientific inquiry skills like these given below:*	
Information Processing Data processing	Develop & apply mental calculation skills. Solve problems & explain reasoning behind the solutions. Sort, classify, sequence, compare, and contrast. Analyse relationships, Locate & collect relevant information.	
Challenge own thinking and knowledge	Individual learner competence	
Problem Solving	Identify & understand a practical problem, plan a procedure, and review solutions.	
Enquiry Skills	Ask relevant questions, define problems, plan action and research, predict outcomes, anticipate consequences, test conclusions, and improve ideas. Researching concepts, test conclusions & improve ideas.	

11. LEARNING OUTCOME 2: NATURE OF SCIENTIFIC KNOWLEDGE:

Key Skills	The learner is able to identify the sources of scientific knowledge and to evaluate knowledge claims, taking cultural and historical contexts into consideration. *It is important for students to understand how scientific knowledge develops.*	
Communication	Speak, listen, understand, and respond effectively, critical reflection/analysis.	
Working With Others	Group work. Awareness & understanding of others needs	
Research skills	Be able to identify strengths and weaknesses in procedures, as well understand the context under which findings were obtained. Judge the limitations and applicability of information.	
Creative Thinking Skills	Apply, generate & extend ideas, suggests hypothesis, apply imagination, find alternative, and innovative procedures.	
Evaluation Skills	Develop judging criteria for procedures and, have confidence in choices. Relate kits to everyday / workplace equipment. Judge value of what is read/heard, and actions. Develop judging criteria for own/others' work/ideas, have confidence in judgements.	

12. LEARNING OUTCOME 3: CONSTRUCTING AND APPLYING SCIENTIFIC KNOWLEDGE

Key Skills	The learner is able to state, explain, interpret, and evaluate scientific and technological facts, concepts, principles, theories, models and laws, and can apply them in everyday contexts.	
Interpreting, data processing	Develop & apply mental calculation skills. Solve problems & explain reasoning behind the solutions. Sort, classify, sequence, compare, and contrast. Analyse relationships	
Draw reasonable conclusions from data	Make reasonable conclusions, test conclusions, & improve ideas.	
Creative Thinking Skills	Apply, generate & extend ideas, suggests hypothesis, apply imagination, find alternative, innovative procedures.	
Evaluation Skills	Relate data to everyday use	

13. LEARNING OUTCOME 4: SCIENCE, TECHNOLOGY, SOCIETY, AND THE ENVIRONMENT

Skill	Science, Technology, Society and the Environment: *This outcome is necessary to help learners to make informed decisions and to have a broader understanding of how science relates to their everyday lives, the environment and to a sustainable future*	
Financial aspects	Value for money. Develop sense of responsibility.	
Enterprise Education	Develop confidence, self-reliance & acceptance of change.	
Education for Sustainable Development	Make informed individual/collective local/global decisions to improve quality of life without damaging the planet for the future.	
Management	**Are you able to organise the programme for all your classes?** What advantages does the programme provide in this regard?	
Safety	Safety of using Somerset kits, and how this informs learners of the need for their safety, and care for the environment	

14. OUTCOMES THAT CANNOT BE DUPLICATED OR DONE THE SAME WAY BY OTHER (NON-COMPUTER-BASED) RESOURCES? E.G., ENJOYMENT, ETC.

15. IN CONCLUSION, IS THE NATURE OF THE SUBJECT WELL-REPRESENTED?

PART E

LEARNERS' EVALUATION OF THE COMPUTER PROGRAMME

GIVE THIS FORM TO A FEW (E.G., 10) LEARNERS AFTER THEY HAVE USED THE PROGRAMME YOU ARE EVALUATING. LEAD THEM THROUGH A DISCUSSION OF THE PROGRAMME. GO THROUGH THIS FORM WITH THEM, EXPLAINING THE MEANING OF EACH ATTRIBUTE. LET EACH ONE OF THEM FILL IN THE FORM, AND COLLECT THE FORMS WHEN THEY FEEL READY TO GIVE A JUDGEMENT.

Name of school		Male		Female		Age		Grade	

16. ALLOCATE A SCORE FOR EACH OF THESE

Attribute	**Marks**											
	0	1	2	3	4	5	6	7	8	9	10	Average
Coverage of content												
a. *Compared to the whole syllabus*												
b. *Compared to the topic covered*												
Feedback												
o *Speed of feedback*												
o *The programme does what I want*												
I understand the plan of the programme												
It is easy to find my way through this programme												
I know the aim of this programme												
I can achieve the aims of this programme												
It promotes learning												
Content is relevant to the syllabus												
It is fun to use the programme												
I would like more learning												
I would like more fun												
I would like more problems to solve												
I have gained more understanding than before												
It is a better way to learn												

17. WHAT EVENTS DO YOU REMEMBER FROM USING THIS PROGRAMME?

____

____

____

18. WHAT INFORMATION DO YOU REMEMBER FROM USING THIS PROGRAMME?

____

____

____

19. WHAT ACTIVITY IN THE PROGRAMME TEACHES YOU MOST?

____

____

____

____

20. WHAT WOULD YOU ADD OR TAKE AWAY FROM THIS PROGRAMME?

____

____

____

____

I give permission fro this information to be used without mention of my names and institution whatsoever. Signed:

…………………………………………………………………………………

A Manual for Zadarh

Johnnie W F Muwanga-Zake

Installing Zadarh

1. Hardware and software requirements for Zadarh
   a. Hardware:

   b. Software:

2. Installation
   Insert the Zadarh CD into the CD ROM of your computer. If your computer is set for 'auotorun', the CD will run automatically and will give you instructions to follow.

   Otherwise, click at the 'Start' icon on the screen of your computer. Search for the CD ROM drive and click open. Look for the 'setup' icon and double click on it. Then follow instructions.

Starting Zadarh

1. Make sure that Zadarh is installed on your computer. A Zadarh icon (Figure 1 below) should appear on the screen if it is installed.

   Figure 1. Zadarh icon on the screen of the monitor
   Or you will see 'Zadarh listed among the programs of your computer if Zadarh is installed on the computer

2. Click on the Zadarh. Figure 2 below will appear

   Figure 2. Opening window of Zadarh
   Note that the pictures inside the window change continuously.

3. This opening window gives you three options:

   a. Start a new game?
   If you click on 'Yes' across this option, a new game will start

   a. Continue to play the last game?
   If you click 'on Yes' across this option, the game which you last played will open at

the point you left it.

a. Exit

If you click on 'Yes' across this option, the opening window will close, and you will not play Zadarh.

Introduction to Zadarh

If you opt to start a new game, Zadarh starts with music together with a series of screens. Examples of these screens appear below:

Screen 1

Screen 2

These introductory screens are followed by instructions about your mission. You can skip the introduction by pressing the button 'Esc' on the key board of your computer.

The entry room

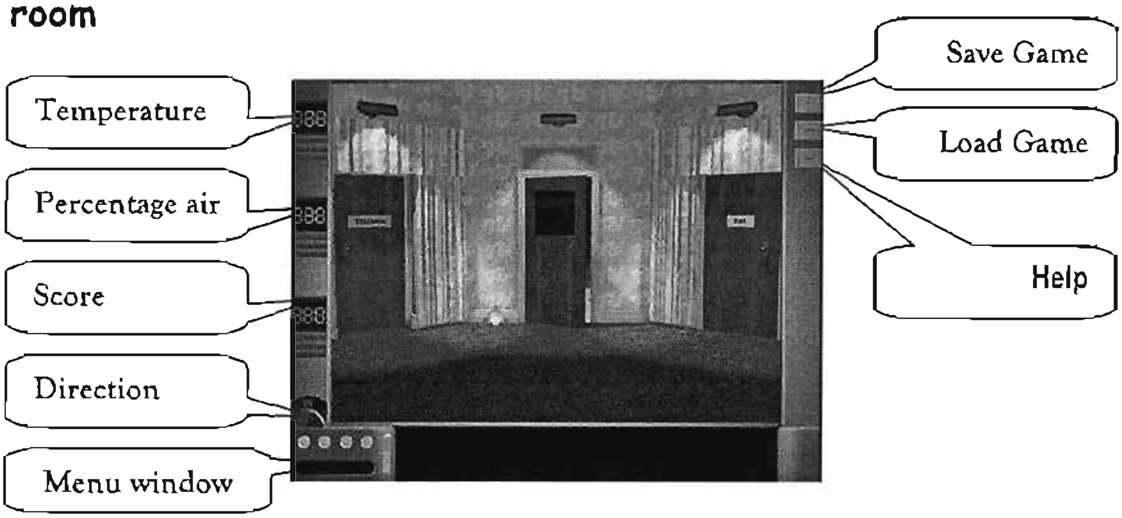


Figure 3. The entry room

This is the first room that you enter when you start anew game.

The menu

Move the cursor around, by moving the mouse. The cursor moves in the same direction as the mouse.

1. *Direction*
    a. There is a compass (see figure 3 above), which indicates to you the direction you are facing
    b. You can only move at right angles, not diagonally

2. *Saving your game (see figure 3 above)*

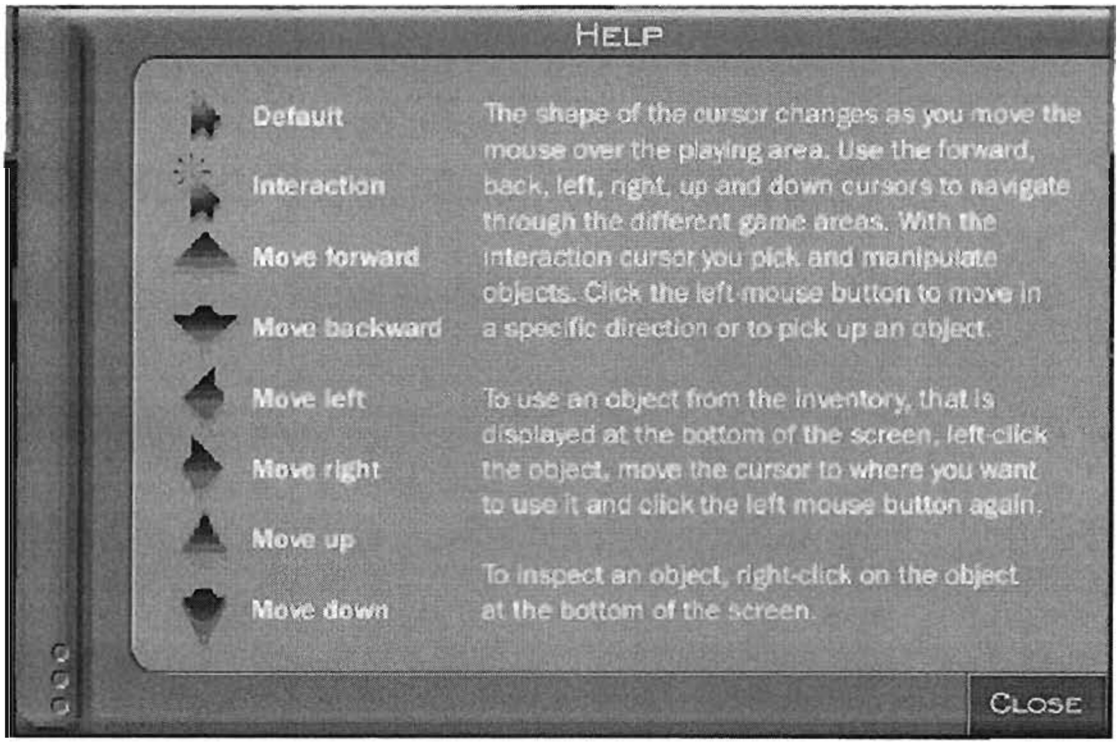
    Click on the 'Save Game' button. A window will appear in the middle of the screen. The window offers you spaces in which you can give a name to your game.

3. *Playing your saved game*

    Click on the 'Load Game' button (see figure 3 above). A window will appear with a list of all saved games. Chose the game you want to replay by clicking on it.

4. *Help button (see figure 3 above)*

    If you click on the 'Help' button, a window shows in the middle of the screen in which you will find explanations of the different shapes of cursors:



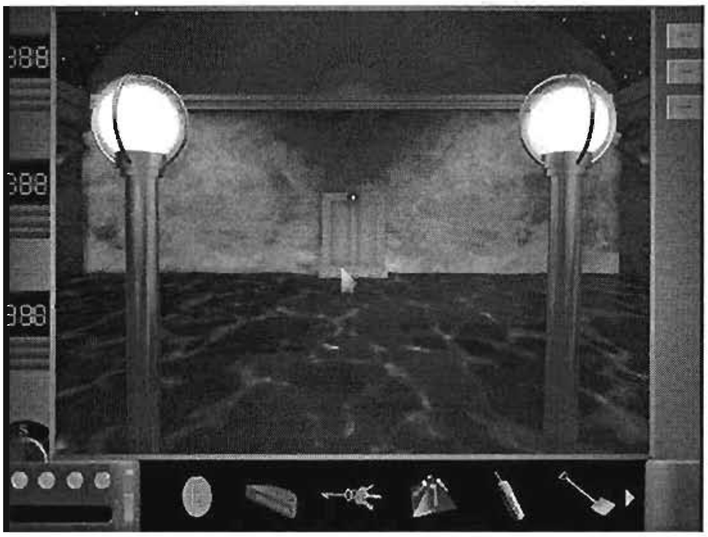
The shape of the cursor changes as you move the mouse over the playing area. Use the forward, back, left, right, up and down cursors to navigate through the different game areas. With the interaction cursor you pick and manipulate objects. Click the left mouse button to move in a specific direction or to pick up an object.

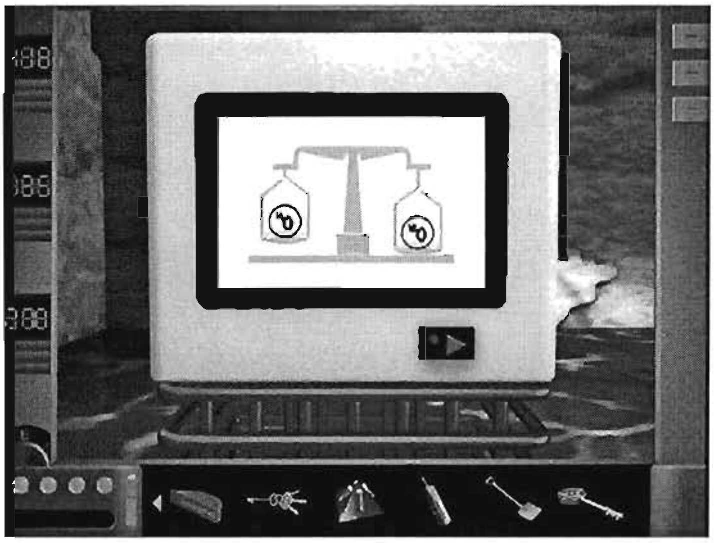
To use an object from the inventory, that is displayed at the bottom of the screen, left-click the object, move the cursor to where you want to use it and click the left mouse button again.

To inspect an object, right-click on the object at the bottom of the screen.

Score

You score every time you solve a problem. You might end up with a similar score by solving different problems. Below are some of the highest scores obtained by learners so far:





Contact for assistance

Please contact me if you get problems or comments about Zadarh

J. W. F. Muwanga-Zake
Office - 031 2603418
Cellular - 0837521534
e-mail jmuwanga@lycos.com

What is Evaluation?
Some Guidelines for School Educators

Johnnie W F Muwanga-Zake

Introduction

There are numerous sources that provide in-depth analyses of evaluations, their designs, methods, and techniques of analysis. This guideline aims at helping obtain an overview of evaluation especially in terms of how you can evaluate teaching resources at your school.

Who Should Carry Out the Evaluation?

Most of our schools do not have specialised evaluators and hiring one is normally beyond school funds. Furthermore, you the educator should be able to make choices from among a wide variety of products or approaches. Making choices involves evaluating those choices. Therefore, it is important that you become an evaluator.

Evaluation and assessment

Evaluation and assessment are both applied in schools. Percival & Ellington (1984: 100) note, they are sometimes in literature used interchangeably. *Check your dictionaries and encyclopaedia.*

However, Fraser (1991: 1) is adamant that although these *concepts are related, they are not synonymous.* For example, Percival & Ellington (1984: 100) as well as Lloyd-Jones, Bray, Johnson & Currie (1986: 1) distinguish between evaluation and assessment. Lloyd-Jones *et al.* point out that evaluation cannot do without assessment.

Assessment

You will find the definition of assessment in many documents from the Department of Education (DoE). In all of them, assessment is the gathering of data about the performance of learners (or even educators).

Evaluation

Evaluation is the collection of information about something in order to make necessary decisions about it. Evaluation can take any form. E.g., formative, summative, etc. The type of evaluation you undertake depends on what you want to learn about what you are evaluating. Don't worry about what type of evaluation you need or are doing - worry about what you need to know, the decisions you need to make, and worry about how you can accurately collect and understand that information.

Key Considerations:

1. For what purposes is the evaluation being done, i.e., what do you want to be able to decide as a result of the evaluation?
2. Who are the audiences for the information from the evaluation: E.g., learners, parents, the DoE, other educators, etc.?
3. What kinds of information are needed to make the decision? E.g., strengths or how what you are evaluating helps.
4. What sources should the information be collected: E.g., learners, staff, etc.
5. How can that information be collected in a reasonable way and time? E.g., questionnaires, interviews, examining documentation, observing learners, conducting focus groups interviews
6. When is the information needed (so, by when must it be collected)?
7. What resources are available to collect the information?
8. What is the objective of what you are evaluating?
9. What values must you take into account? E.g., gender biasness, racism, language, etc.
10. Does appearance matter?
11. Is the product needed and by who?

Ethics: Informed Consent from participants

An evaluation might require you to obtain the views of fellow staff members or of learners. You should first gain the consent of participants if you plan to include in your evaluation, the focus and reporting on their personal information. They should understand what you're doing with them in the evaluation and how any information associated with them will be reported. You should clearly convey terms of confidentiality regarding access to evaluation results. They should have the right to participate or not. Have participants review and sign an informed consent form.

Overview of Methods to Collect Information

The overall goal in selecting evaluation method(s) is to get the most useful information to make choices in the most cost-effective and realistic fashion. The following table provides an overview of the major methods used for collecting data during evaluations.

Method	Overall Purpose	Advantages	Challenges
Questionnaires, surveys, checklists	When need to quickly and/or easily get lots of information from people in a non threatening way	-Can complete anonymously inexpensive to administer-easy to compare and analyse administer to many people -can get lots of data-many sample questionnaires already exist	-Might not get careful feedback wording can bias client's responses -are impersonal in surveys, may need sampling expert - doesn't get full story
Interviews	When want to fully understand someone's impressions or experiences, or learn more about their answers to questionnaires	-Get full range and depth of information -develops relationship with client -can be flexible with client	-Can take much time can be hard to analyse and compare can be costly interviewer can bias client's responses
Documentation review	When want impression of how a product operates.	-Comprehensive -information already exists few biases about information	-Often takes much time -info may be incomplete -need to be quite clear about what looking for
Observation	To gather accurate information about how a program actually operates, particularly about processes	-View operations of a program as they are actually occurring can adapt to events as they occur	-Can be difficult to interpret seen behaviours-can be complex to categorize observations -can influence behaviours of program participants -can be expensive
Focus groups	Explore a topic in depth through group discussion: E.g., about reactions to an experience or suggestion, understanding common complaints, etc.; useful in evaluation and marketing	-Quickly and reliably get common impressions can be efficient way to get much range and depth of information in short time can convey key information about programs	-Can be hard to analyse responses -need good facilitator for safety and closure -difficult to schedule 6-8 people together
Case studies	To fully understand or depict client's experiences in a program, and conduct comprehensive examination through cross comparison of cases	-Fully depicts client's experience in program input, process and results -powerful means to portray program to outsiders	-Usually quite time consuming to collect, organize and describe represents depth of information, rather than breadth

Analysing data

Always start with your evaluation goals:

When analysing data (whether from questionnaires, interviews, focus groups, or whatever), always start from reviewing your evaluation goals, i.e., the reason you undertook the evaluation in the first place. This will help you organize your data and focus your analysis.

Basic analysis of "quantitative" data

- Make copies of your data and store the master copy away. Use the copy for making edits, cutting and pasting, etc.
- Read through all the data.
- For written or recorded responses, attempt to identify patterns, or associations and causal relationships in the themes: E.g., all people who attended programs in the evening had similar concerns, most people came from the same geographic area, most people were in the same salary range, what processes or events respondents experience during the program, etc.
- Organize comments into similar categories: E.g., concerns, suggestions, strengths, weaknesses, similar experiences, program inputs, recommendations, outputs, outcome indicators, etc.

- Tabulate the information, i.e., add up the number of ratings, rankings, yes's, no's for each question.
- For ratings and rankings, consider computing a mean, or average, for each question. For example, "For question #1, the average ranking was 2.4". This is more meaningful than indicating: E.g., how many respondents ranked 1, 2, or 3.
- Consider conveying the range of answers: E.g., 20 people ranked "1", 30 ranked "2", and 20 people ranked "3".

Interpreting Information:

Attempt to put the information in perspective: E.g., compare results to what you expected; any common standards for your services; original program goals (especially if you're conducting a program evaluation); description of the program's experiences, strengths, weaknesses, etc.

Writing the report

1. The level and scope of content depends on to whom the report is intended.
2. Record conclusions and recommendations in a report document, and associate interpretations to justify your conclusions or recommendations.
3. Consider recommendations to contribute towards improving the product.
4. Remember to document the evaluation plans and activities, which can be referenced when a similar evaluation is needed in the future.
5. Purpose of the report - what decisions are being aided by the findings of the evaluation
6. Background what is being evaluated

The report

1. Description/History
2. Overall Evaluation Goals (E.g., what questions are being answered by the evaluation)
3. Methodology
   a. Types of data/information that were collected
   b. How data/information were collected (what instruments were used, etc.)
   c. How data/information were analysed
4. Limitations of the evaluation (e. g, cautions about findings/conclusions and how to use the findings/conclusions, etc.)
5. Interpretations and Conclusions (from analysis of the data/information)
6. Recommendations (regarding the decisions that must be made about the product/service/program)
7. Appendices: content of the appendices depends on the goals of the evaluation report: E.g.:
   a. Instruments used to collect data/information
   b. Data: E.g., in tabular format, etc.
   c. Testimonials, comments made by users of the product/service/program
8. Any related literature

EXERCISES

1. Make a list of products at your school, which need to be evaluated
2. One of the most common products at school is a textbook. Take one of your science textbooks (include biology) and evaluate it.
   a. Start by choosing the aspects of the textbook you will evaluate. For example, the quality of diagrams, completeness of content, cultural bias, correctness of content, etc.
   b. List the aspects you have chosen. Ask a colleague whether all the important aspects of the textbook have been catered for.
   c. Evaluate the book
   d. Make an evaluation report
3. Imagine that your school wants to buy a computer programme for teaching science. How would you go about choosing the best one in the market.

CONTACT FOR ASSISTANCE

J. W. F. Muwanga-Zake: Office – 031 2603418 Cellular – 0837521534 e-mail jmuwanga@lycos.com