UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Bridging the Gap

Optimising a Feedback System for Monitoring Learner Performance

by
Elizabeth Archer
2010

Presented in partial fulfilment of the requirements for the degree

DOCTORATE OF EDUCATION

(ASSESSMENT AND QUALITY ASSURANCE)

Faculty of Education
University of Pretoria

Supervisors

Professor SJ Howie
Professor B Smit

This thesis is dedicated to my beautiful, loving wife, Catherine.

We made it, babes!

Acknowledgements

Many people directly and indirectly contributed to the process of creating this thesis. I cannot thank all of you individually, but thank you for helping me on this journey. I would however like to thank some very special people:

- All the teachers, HoDs and principals who gave so freely of their time and expertise to help build this feedback system and who contributed to this thesis. Thank you for sharing your expertise, insights and passion.

- Sarah, I really enjoyed our thinking sessions and the way you challenged me intellectually. Thank you for so generously providing me with the time and support I needed for my thesis. The opportunities you have created for me are deeply appreciated.

- Brigitte, it was your mentorship and guidance which first stimulated my interest in research nearly a decade ago. Thank you for opening this door to the world of rigorous enquiry for me.

- Vanessa, you are an amazing friend. Thank you for always listening, supporting and providing advice. You are a technical advisor *par excellence*. May God bless you richly.

- Prof Tjeerd Plomp, thank you for being a critical friend and true mentor. I was blessed to benefit from your guidance and wisdom. I hope my retirement one day will be a bit more quiet than yours ☺.

- Prof Roel Bosker, thank you for your guidance and support as well as the learning opportunities you made possible at GION at the University of Groningen.

- My fellow PhD sufferers, Kim and Lisa. Our impromptu debriefing, support, consultation and therapy sessions always nourished my emotional well-being. Thank you for being there and sharing the load.

- The South African Netherlands Research Programme on Alternatives in Development (SANPAD), who generously funded the SAMP project. Thank you also for the opportunity to attend the SANPAD Advanced Research Capacity Initiative, which stimulated my growth as a researcher. The collaboration with our Dutch partners has been invaluable.

- The National Research Foundation (NRF), thank you for providing the funding for this thesis and initial funding for the PIPSSA project. Without funders such as the NRF post-graduate studies would be out of reach for most South Africans.

- Most of all I want to say thank you Catherine. Without your love and support this would not have been possible. Thank you for believing in me, supporting me, looking after me and putting up with me. I love you always.

Abstract

Globally, a wealth of educational data has been collected on learner performance in a bid to improve and monitor the quality of education. Unfortunately, the data seem to have had only limited influence on learning and teaching in classrooms. This thesis aimed to bridge this gap between the availability of learner performance data and their use in informing planning and action in schools. A design research approach was used to optimise the feedback system for the South African Monitoring system for Primary schools (SAMP).

Design research aims to produce both an intervention to address a complex real-world challenge and to develop design guidelines to support other designers faced with similar challenges in their own context. In this research, the process of developing and improving the feedback system was also used to examine ways of facilitating the use of the feedback. Multiple cycles of design, implementation and evaluation of four different prototypes of the feedback system were conducted, employing evaluations from both experts (e.g. Dutch and South African academics, research and educational psychologists, instrument designers and teacher trainers) as well as school users (teachers, principals and HoDs).

Mixed methods were employed throughout the study, with different sub-samples of school users sampled from the population of 22 schools (English, Afrikaans and Sepedi) in the Tshwane region participating in SAMP. The various research cycles incorporated interviews, observations, journals, questionnaires, the Delphi technique and expert evaluations to examine not only data-use, but also aspects such as problem-solving, planning, data-literacy and attitudes towards evidence-based practice in the schools. Data was analysed using Rasch Modelling, descriptive statistics and computer-aided qualitative data analysis.

The study showed that an effective feedback system facilitates appropriate use through a gradual process of enlightenment, is flexible and responsive to user inputs, values collaboration and includes instrument, reporting and

support components in its design. An optimum feedback system also positively influences school feedback and monitoring culture by providing opportunities for positive experiences with feedback and increasing data-literacy. This improves the chances of feedback being used for planning, decision-making and action in the schools. An effective feedback system must also offer a comprehensive package to accommodate different users, with various levels of data sophistication, functioning in diverse contexts. The research also showed that an effective feedback system mediates thinking about educational instruction and curriculum and can therefore be a potent change agent. Use of clear, simple, intuitive data presentation in the feedback system allows for experiential learning to increase user data-literacy.

The design research approach employed in this study offers an appropriate and powerful approach to adapting, developing and optimising a feedback system. User involvement in design research ensures greater contextualisation and familiarity with the system, while engendering trust and a greater sense of ownership, all of which increase the receptiveness and responsiveness of users to feedback. Finally, the research also contributed design guidelines for other developers of feedback systems, an integrated conceptual framework for use of monitoring feedback and a functioning feedback system employed by 22 schools in the Tshwane region.

Keywords: Feedback use; Data use; Data-literacy; Statistical-literacy; Learner performance monitoring; Feedback system; School Performance Feedback System (SPFS); Evidence-based practice; Design Research; Enlightenment

TABLE OF CONTENTS

CHAPTER THREE
CONTEXTUALISATION, LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK - 54 -

CHAPTER FOUR
OVERVIEW OF THE RESEARCH DESIGN - 92 -

CHAPTER FIVE
PRELIMINARY PHASE: PROBLEM IDENTIFICATION, NEEDS AND CONTEXT ANALYSIS - 123 -

CHAPTER SIX
PROTOTYPING PHASE: ESTABLISHING CONDITIONS FOR USE (CYCLE 1-2) - 165 -

LIST OF FIGURES

TITLE **PAGE**

LIST OF TABLES

TITLE **PAGE**

LIST OF ABBREVIATIONS

450 Support	Forms to be completed when child is identified as at risk of failure and submitted to the DoE, must be accompanied by an individualised support program.
ABC+	Attitudinal/Behavioural/Cognitive Indicators, plus Context
ABEL	Assessment for Better Learning
AIDS	Acquired Immune Deficiency Syndrome
ALIS	A Level Information System
ARB	Assessment Resource Bank
ASPECTS	Assessment Profile on Entry for Children and Toddlers
asTTle	assessment Tools for Teaching and learning
AtoL	Assess to Learn
CAQDAS	Computer Aided Qualitative Data Analysis
CASS	Continuous assessment
CD	Compact Disc
CEA	Centre for Evaluation & Assessment
CEM	Centre for Evaluation and Monitoring
CREATE	Consortium for Research on Education, Access, Transitions & Equity
DIF	Differential Item Functioning
DoE	Department of Education (South Africa)
DVD	Digital Video Disc
ERO	Education Review Office
GDE	Gauteng Department of Education
GDP	Gross Domestic Product
GET	General Education and Training
HIV	Human Immune Deficiency Virus
HoD	Head of Department

HSRC	Human Science Research Council
ICT	Information and Computer Technology
InCAS	Interactive Computer Assessment System
IQMS	Integrated Quality Management System
LEAs	Local Education Authorities
LOLT	Language of Learning and Teaching
MidYIS	Middle Years Information System
MLA	Monitoring Learning Achievement
NEIMS	National Education Infrastructure Management System
NEMP	National Education Monitoring Project
NQF	National Qualifications Framework
NRF	National Research Foundation
NZ MoE	New Zealand Ministry of Education
NZ	New Zealand
OBE	Outcomes Based Education
OECD	Organization for Economic Co-operation and Development
OFSTED	Office for Standards in Education
OT	Occupational Therapist
PAT	Progressive Achievement Tests
PD	Professional Development
PIPS	Performance Indicators in Primary Schools
PIPSSA	Performance Indicators in Primary Schools in South Africa
PIRLS	Progress in International Reading Literacy Survey
PISA	Programme of International Student Assessment
RSA	Republic of South Africa

SACMEQ	Southern African Consortium for Monitoring Educational Quality
SAM	School Analysis Model
SAMP	South African Monitoring system for Primary schools
SANPAD	South African Netherlands Research Programme on Alternatives in Development
SAQA	South African Qualifications Authority
SASSIS	South African Secondary School Information System
SBST	School Based Support Team
SGB	School Governing Body
SIOP	Sheltered Instruction Observation Protocol
SITES	Second Information Technology in Education Study
SOLO	Structure of Observed Learning Outcomes
SPFS	School Performance Feedback System
STAR	Supplementary Test of Achievement in Reading
TIMSS	Third/Trends in International Mathematics and Science Study
UK	United Kingdom
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
UP	University of Pretoria
USA	United States of America
YELLIS	Year 11 Information System
ZEBO	*Zelf Evaluatie in het Basis Onderwijs*[1]

[1] English translation: *Self-Evaluation in Primary Education*

CHAPTER ONE

Introduction and Overview

> …while volumes of data are extruded about and from schools, teaching continues without the benefits of such data. There is still a philosophy that assumes teachers know how and what data to collect to best enhance learning, and many of these assumptions are based on folk philosophies, poor measurement, and shaky data. We still teach in a manner we did 150 years ago … (Hattie, 2005, p. 11)

Globally a wealth of educational data has been collected on learner performance, but while this has had a major impact on systemic and curriculum reform (Howie & Plomp, 2005) it seems to have had only limited influence on learning and teaching in classrooms (Fullan, 2006; Hattie, 2005; Patton, 1991, 1997; Visscher & Coe, 2002). It is a global phenomenon that even when school data are available, schools often do not use them or are unable to interpret them correctly and consequently, use the data for improvement of quality of education in the schools (Schildkamp & Teddlie, 2008; Wohlstetter, Datnow, & Park, 2008). This can also be observed in the Republic of South Africa (RSA), where despite the available educational data, formal monitoring and use of data to inform planning, teaching and learning in schools is still limited (Department of Education, 2002b, 2006a, 2006b).

This thesis aims to bridge this gap between the availability of learner performance data and their use in informing planning and action in schools. The thesis employed a design research approach to determine design guidelines and principles[2] to facilitate the use of feedback[3] from learner performance monitoring and enhance the use of data at school level. Design research has two purposes:

[2] Van den Akker (1999) describes design principles as the major knowledge gained from design research. Design principles can be either substantive or methodologically orientated and aim to act as guidelines to other designers faced with a similar problem to support their design tasks.

[3] The use of feedback in this case is not seen as only the use of monitoring information, but also any process use of the skills and knowledge gained by participants during the monitoring process. This conceptual use may include changes in how assessment is viewed, different approaches to problem-solving, and greater confidence in interacting with data, etc. The issue of use of feedback is explored in depth under the conceptual framework in Chapter 3.

1. To design, develop or optimise a product or intervention for a complex real-world problem
2. To develop design guidelines to contribute to the body of knowledge on a phenomenon and support other designers faced with similar challenges in their own context

   (De Villiers, 2005; Plomp, 2009; Van den Akker, 1999)

For the purpose of this study, an existing feedback system known as the South African Monitoring system for Primary schools (SAMP) was optimised (see Section 1.2). The process of developing and improving the feedback system was used to examine ways of facilitating the use of the feedback in participating schools. Design guidelines were produced to facilitate use of learner performance feedback systems. The study examined not only data-use, but also implicit changes in aspects such as problem-solving, planning, data-literacy (see Section 1.1 for definition) and attitudes towards evidence-based practice in the schools.

This chapter first provides the definitions of the terms as used in this study, given the plethora of terms in this field (Section 1.1). Thereafter, the SAMP project that is the focus of this study is discussed (Section 1.2). This is followed by an illustration of the study's importance and value, with the problem presented in context and the rationale given (Section 1.3). The general research questions are introduced in Section 1.4 to be operationalised in Chapters 3-8, based on the literature review, conceptual framework and research procedures. Next, the research methods and design are briefly discussed to frame the empirical component of the inquiry (Section 1.5). This is followed by a discussion of the researcher's role and positioning in this research (Section 1.6). The chapter concludes with an outline of the remaining chapters that comprise this thesis (see Section1.7).

1.1 Definition of Terms

The literature for this study employs various terms such as monitoring, evaluation, assessment, quality assurance, feedback, data, data-literacy, feedback systems and feedback facilitators, all of which refer to measurement

and the use of information. The definitions of these terms are constantly being debated, but working definitions are presented here to provide a common understanding of use in this thesis.

Traditionally, ***evaluation*** is seen as systematic gathering of information that results in judgement (Beeby in Husėn & Tuijnman, 1994, p. 1; Scheerens, Glas, & Thomas, 2003). The Joint Committee on Standards for Educational Evaluation (1994) based their definition of evaluation on the work of Scriven (1967, p. 39) who states that evaluation is "[t]he systematic investigation of worth or merit of an object". For the purpose of this study, these two sources are drawn on to define evaluation as the systematic gathering of data to investigate the worth or merit of the feedback system.

Monitoring involves the tracking of change over time that demands systematic and regular procedures for the collection of data at multiple points in time for decision-making (Husėn & Tuijnman, 1994; Nuttal, 1994). Scheerens et al. (2003) see monitoring as a type of evaluation that calls for ongoing information gathering focused on description rather than valuing and judgement. For the purpose of this study, monitoring is defined as a type of evaluation that requires a systematic collection of data at multiple points in time for the purpose of decision-making. Monitoring focuses on providing information that participants can use to make decisions and apply value and judgement. The aim of these decisions is to improve or enhance education.

Quality assurance is described by Scheerens et al. (2003) as the purpose of monitoring in education. In this thesis, it relates to monitoring and evaluating the quality of teaching, learning, planning and action taken in schools to improve learning.

Assessment is a measurement activity that usually refers to the measurement of learner achievement in the context of education (Joint Committee on Standards for Educational Evaluation, 1994). Hattie and Jaeger (1998, p. 116) note that the difficulty with assessment is that knowledge acquisition is mostly assessed while "…other functions of learning such as deep understanding, efficient intuitive use, acquiring multiple flexible

strategies, adaptive action control, and achievement motivation" are ignored. Assessments are often employed for the purposes of monitoring and evaluation. In this study learners' phonics, reading and mathematics skills are assessed using the South African Monitoring system for Primary schools (SAMP). Therefore, assessment refers here to the baseline and follow-up measurements conducted with learners in SAMP. This forms the basis of the data used in the feedback system.

In the context of education, ***feedback*** relates to returning data gathered through monitoring or evaluation to schools, learners or parents. Hattie (2005) states that feedback can be complex, in particular when considering how best to present data in a comprehensible and accessible form to schools. The aim of attending to data presentation and feedback mechanisms is to facilitate use of the feedback to lead to improvement in schools.

Data[4] consist of discrete, objective evidence collected by qualitative or quantitative methods. Data could be drawn from observation, anecdote, opinions, or figures (such as averages). To qualify as data, systematic inquiry, organisation and analysis that incorporate various views is required (Earl & Katz, 2006). Interpreted data are known as 'information' and have meaning, relevance and purpose (Schildkamp & Kuiper, 2009). In this study, learner performance data were generated through the SAMP assessments. Evaluation data from expert evaluators, teachers, heads of department (HoDs) and principals were also generated through questionnaires, observations, interviews, evaluation reports as well as the Delphi technique.

[4] Data is a Latin plural (singular datum), but is generally defined as uncountable and therefore may take the singular form of the verb.

Data-literacy in this thesis refers to a school's ability to interact actively with data, interpret and apply them. According to Earl and Katz (2006, pp. 19-20) a data-literate school leader is able to:

- think about the purpose for which the data were generated
- recognise sound and unsound data
- be knowledgeable about statistical and measuring concepts
- recognise that there are several types of data
- make interpretation paramount
- pay attention to how data are reported and for which audience

Feedback systems are information systems that focus on generating and providing data to users. In the educational context a specific brand of feedback system can be identified, namely School Performance Feedback Systems (SPFSs), which are "external to schools that provide them with confidential information on their performance and functioning as a basis for school self-evaluation" (Visscher & Coe, 2002, p. xi). The SAMP project on which this study is based is classified as a School Performance Feedback System (SPFS).

In this thesis, a ***feedback facilitator*** is conceptualised as the person responsible for coordinating the monitoring process, compiling reports, and providing feedback to participants, as well as supporting interpretation and encouraging action based on the monitoring and feedback process. In this case, I, as researcher also fulfil the role of monitoring facilitator, but this role may eventually shift to the district office of the Department of Education (DoE) or even individual school coordinators, as the SAMP system develops and becomes more widespread. Sections 1.6 and 4.3.1 provide further discussion of the dual role of facilitator and researcher in this study.

1.2 The SAMP Project

Data for educational improvement purposes can be generated in multiple ways and be managed either externally or internally. Data can also be aggregated and related to different levels of education, including the systemic,

school and classroom levels. Different types of data such as those on learner performance, teacher performance and management functioning can inform educational improvement. This research focuses on optimising a feedback system for a school-based monitoring system, SAMP, which is currently facilitated externally by the Centre for Evaluation and Assessment (CEA). The system produces learner performance data that are also aggregated to school level. The data from SAMP are employed to inform individual learner intervention, classroom practice, and school level planning and action.

SAMP focuses on the entry level to primary education known as Grade 1 (usually five to seven years of age). This forms part of the Foundation Phase of Education, designed to establish basic literacy, numeracy and phonics skills so that learners can progress to learn more independently in the later phases of schooling. V. Greaney (personal communication with SJ Howie, April 18, 2006) stresses the importance of the establishment of foundational skills, since learners who are not able to read at 12 years of age are unlikely to learn to read later in life. Poor foundational skills have also been noted as a large contributory factor to poor learner performance in secondary education by the DoE (Department of Education, 2006b, 2006d)

SAMP produces reliable and valid data for the South African context across the three languages in which it is currently employed, namely English, Afrikaans and Sepedi (for a fuller discussion of the quality of the instrument and current functioning, see Chapter 2). SAMP therefore provided relevant and appropriate data to use in the feedback system and offered the opportunity to study the use of the feedback.

SAMP is a research project originally funded by the National Research Foundation (NRF) to develop monitoring systems at primary and secondary school levels. It was initiated in 2003 in collaboration with the Centre for Evaluation and Monitoring (CEM) at the University of Durham in the United Kingdom (UK). The collaboration aimed to develop a monitoring system suitable for South African education and to develop expertise and resources in the field of learner performance monitoring. The CEM developed a suite of value-added assessments in conjunction with teachers and Local Education

Authorities (LEAs) staff in the UK. The suite employs a combination of objective assessments and teacher ratings to provide valuable information about each learner to schools and teachers (Tymms & Albone, 2002).

The CEA decided to focus on the CEM's assessments for the beginning of primary school (Performance Indicators in Primary Schools or PIPS) and the beginning of secondary school (Middle Years Information System or MidYIS). The entry phases for primary and secondary school were identified as periods when South African schools have limited information about the levels of knowledge and skills of their learners, since learners are often from a large and diverse feeder area (Scherman, Archer, Howie, & Lopez, 2006). The primary school counterpart of the CEM suite in South Africa is known as SAMP, while the South African secondary school component is known as the South African Secondary School Information System (SASSIS).

The PIPS assessment was originally developed in England by the CEM in 1994 (Tymms, Merrell, & Jones, 2004) with the aim of providing not only data on current attainment of learners on curriculum aligned subtests, but also providing predictive data on future mathematics and reading performance (Tymms, Merrell, & Henderson, 2000). The PIPS monitoring system utilises two assessments: a baseline assessment implemented at the beginning of the year and a follow-up assessment administered at the end of the year. This means that both the learners' current level of performance and the difference between the baseline and follow-up performance are reported (for a description of value-added measures, see Section 2.2.1). In 2002, more than a million primary school learners across the globe participated in PIPS (Tymms & Coe, 2003). Figure 1.1 outlines the history of the SAMP project and instruments.

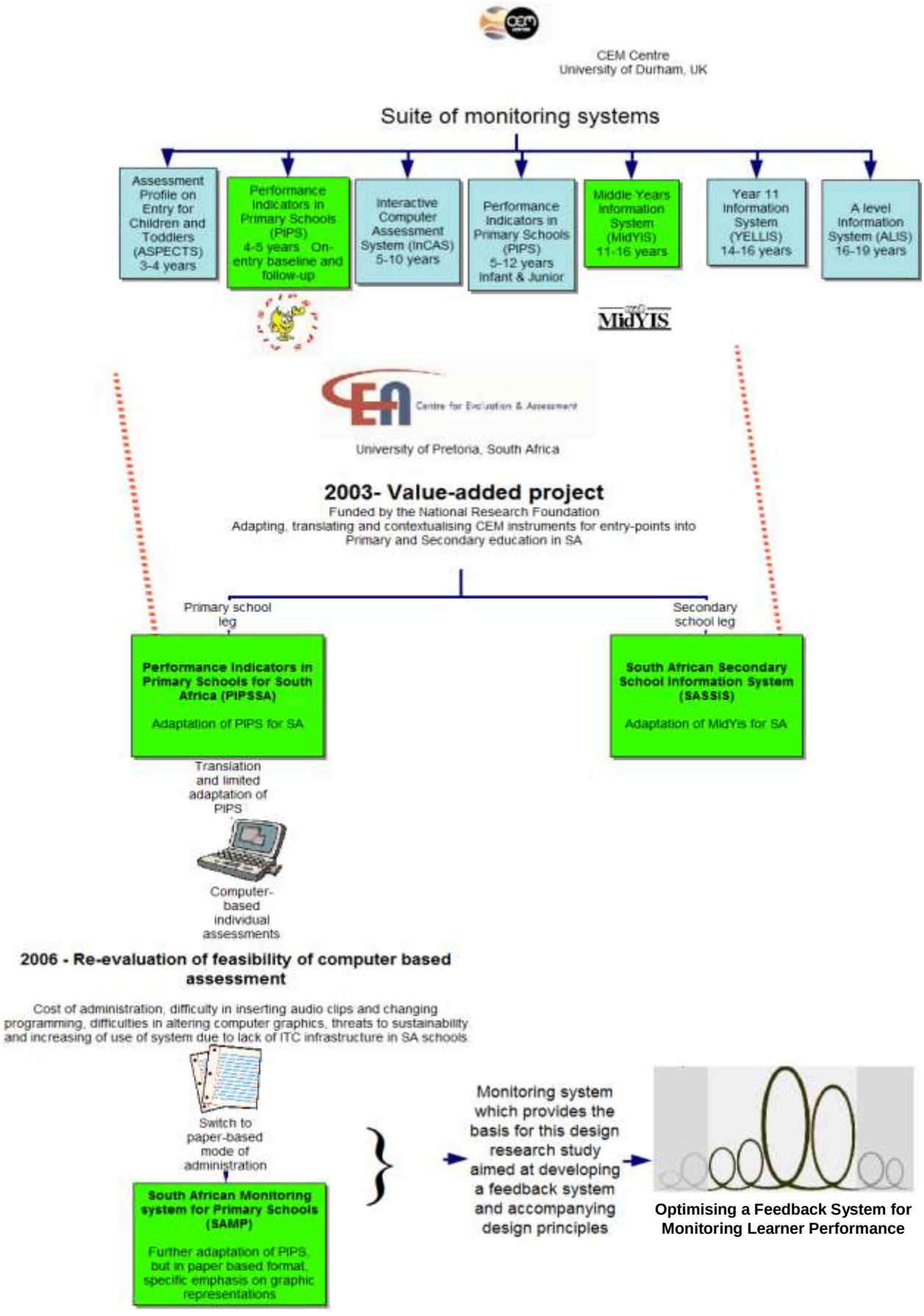


Figure 1.1: History of the SAMP project

The translation and adaptation process of PIPS was initiated in 2003 and originally centred on the computer-based mode of administration of the PIPS instrument, known as Performance Indicators in Primary Schools for South Africa (PIPSSA). During the first three years, the assessment was translated into Afrikaans and Sepedi[5] to determine the feasibility of using the translated PIPS computer-based assessment with the original graphics and instructions in the South African context. The items were also assessed in terms of the national curriculum. The computer-based assessment was recorded in Afrikaans, Sepedi and South African English. These three languages represent the dominant Languages of Learning and Teaching (LOLT) in the Tshwane region where fieldwork took place (H. Julies, DoE, personal communication, July 16, 2010).

The research for this thesis commenced in 2006, the first part of the year being dedicated to consolidation and use of the computer-based assessment data from the previous three years (2003-2005). The project had reached a critical point where the use of the computer-based assessment was not sustainable and limited the expansion or further adaptation of the project. To address this, the next three years of adaptation and development of the PIPS instrument to the South African context focused on paper-based assessment (Archer, 2006a). The project name was also changed to the South African Monitoring system for Primary schools (SAMP) at this stage. The switch to a paper-based assessment immediately increased the feasibility and sustainability of the project in the country's schools as the necessary information and computer technology (ICT) was not yet in place to allow for sustainable computer-based assessment (Department of Education, 2003b; Gauteng Department of Education, 2005). (See Chapter 2 for a full discussion)

The contextualisation process for SAMP was extensive, with various phases of discussion, adaptation and further development of the items incorporating expert appraisal (including Foundation Phase academics, instrument developers, educational and research psychologists and teachers) as well as

[5] A Zulu translation also took place but development was not continued and the SAMP assessment is currently only administered in three languages: English, Afrikaans and Sepedi.

statistical analyses (Archer, 2006a). This process is discussed in full in Chapter 2.

The adaptation process culminated in the SAMP instrument that was employed in 2008 to assess 1,535 learners in English, Sepedi and Afrikaans. The instrument assessed the handwriting, phonics, early reading and early mathematics skills of the Grade 1 learner. Given the South African context, an English Additional Language instrument for non first-language English speaking learners was also developed to assess proficiency in English. The SAMP assessment in its paper-based format consists of a number of subtests, constituting four scales (Early Phonics, Early Reading, Early Mathematics and Handwriting) and the English Additional Language Assessment (as illustrated in Figure 1.2).

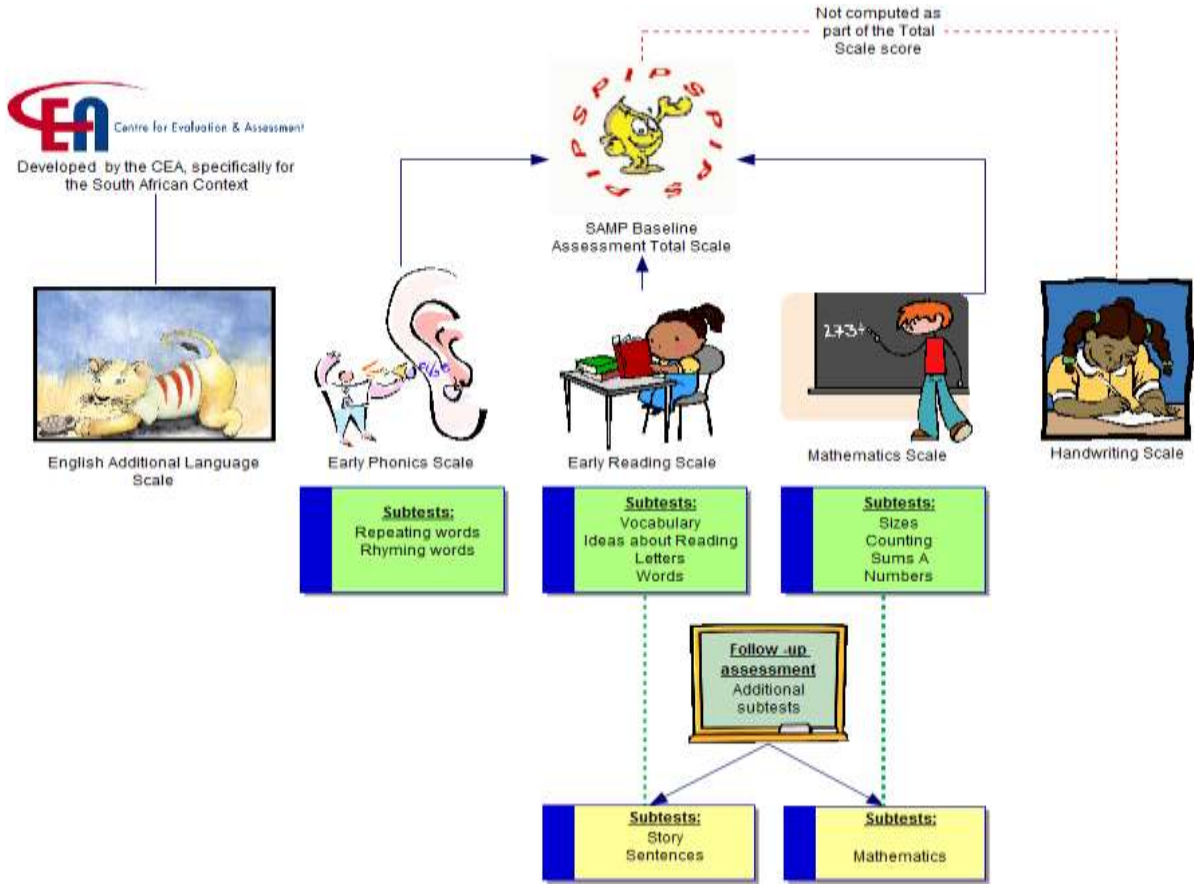


Figure 1.2: SAMP Baseline and follow-up assessment

The baseline assessment consists of 12 subtests[6]; while the follow-up assessment contributes a further three subtests, as indicated in Figure 1.2. The results from the follow-up and baseline assessment are compared and reported on to the schools in terms of gains or losses made.

As a result of the rigorous development and adaptation process (discussed in Chapter 2), the SAMP instruments can be used with confidence as valid and reliable instruments in the South African context. It is therefore appropriate to examine the use of a feedback system based on the data generated by SAMP. In the next section the problem statement and rationale for this study are explored.

1.3 Problem Statement and Rationale

> … [South Africa] is a country with natural wealth and many cultures. It is also notorious for the Apartheid (sic) policies that have left a lasting impression on the education system in the country. Evidence of this [lasting impression] lies in the appalling conditions in many schools across the country, and these conditions exist primarily in previously so-called African, coloured and Indian schools. South Africa, since the first democratic elections in 1994, has embarked on a substantial reform effort in many areas including education. (Howie, 2002, p. 9)

Education is a major concern for the South African government, which had invested 5.8% of the gross domestic product (GDP) in this sector between 1995 and 2003 (National Treasury Republic of South Africa, 2005). In 2006 this investment dropped slightly to 5.4% of GDP, representing 17.6% of the total government expenditure (World Bank, 2008). On average, other upper middle income countries and Sub-Saharan Africa countries spent 4.1% and 4.2% of their GDP respectively on Education in 2006 (World Bank, 2008). The National Treasury of RSA (2005) noted that there had been an enormous growth in enrolment figures for primary and secondary schools (8.1 million in 1985 to 12.0 million in 2004). Population growth and immigration had no doubt contributed to these figures, as the net enrolment rate in primary education dropped from 90% in 1991 to 88% in 2006 (World Bank, 2008). In

[6] The Handwriting and English additional language assessments consist of only one subtest each, which are also used as the scale scores.

2006, South Africa still had nearly 470,000 primary school children aged children who did not attend school (World Bank, 2008).

Despite the significant funding and increase in enrolment for education in South Africa, the quality of education remained a concern (Taylor, Muller, & Vinjevold, 2003). Nowhere was the shortcoming of education provision more apparent than in the low learner performance, especially in subjects such as Reading, Mathematics and Science. This low learner performance was clearly illustrated through South Africa's performance in international studies such as the Trends in Mathematics and Science Study (TIMSS) 2003 (Martin, Mullis, Gonzalez, & Chrostowski, 2004) and the Progress in International Reading Literacy Study (PIRLS) 2006 (Howie, et al., 2008). The concerns about South African education are further highlighted in national studies such the Grade 3 and 6 National Systemic Evaluations (Department of Education, 2002b).

It may be that this poor performance is a legacy of the apartheid education system, however, there is evidence of an international trend wherein increased investment in education is not necessarily associated with improvement in education (Cassassus, 2001; Hayward & Hedge, 2005). Hattie (2005, p. 12) notes that in the United States of America (USA) "…there is not a lot of evidence that the massive increases in state/federal monies have made a difference to the quality of teaching and learning." Hattie (2005) goes on to argue that, though USA spending on education had increased in the previous 40 years[7], the achievement curve remained constant over the same period of time. RSA is challenged with redressing the neglect of large portions of the education system during the apartheid era, but it is clear that this large investment in formal education alone is not improving the quality of learning in schools. This re-asserts the need to combine educational investment with appropriate monitoring at primary school level to improve the quality of teaching and learning.

[7] This increase in USA education spending remains, even if controlled for inflation. The figures can also not merely be downplayed to larger enrolment figures, as this increase remains, even if spending is viewed in terms of investment per learner and when costs for building new schools are factored out (Hood, 1990).

In RSA, educational data are collected through international comparative educational studies such as PIRLS (Howie, et al., 2008) and TIMSS 1995 (Howie, 1997), 1999 (Howie, 2001) and 2003 (Martin, et al., 2004). Systems level data are generated through systemic evaluations, which mirror the poor international performance. Poor performance is noted in both the Grade 3 and Grade 6 National Systemic Evaluation Reports (Department of Education, 2002b, 2006a, 2006b). School level monitoring is also mandated as part of the Internal Quality Management system of schools (IQMS) (Education Labour Relations Council, 2003). The mere availability of these data alone cannot improve learner performance, as the data also needs to be appropriately returned to schools and employed by them for planning, decision-making and action.

Feedback of data on learner performance to teachers and principals has long been regarded as generally enhancing performances of schools and learners (Coe, 2002). Kluger and DeNisis (1996) conducted a meta-analysis of 131 studies where feedback interventions were employed to bring about improvement in schools and they found that the average effect was moderately positive. However, more than 38% of the studies showed a negative effect of feedback of data and the mode of the effect size was zero. This research showed that, although feedback systems could have a positive effect at a school level, they do not have a consistently positive effect in all schools. Although this is the case, many countries around the world have turned to research and monitoring to improve the quality of education, teaching and learning (Jansen, 2001). Coe (2002) and Hattie (2005), amongst other authors, suggest that a great divide exists between monitoring and use of feedback of monitoring data in education to bring about improvement through praxis. Hattie (2005, p. 11) also notes that while there is an abundance of data collected and generated on schools the crucial question remains: “How can we return it to schools?”. In other words, how can the data be suitably fed back into schools to facilitate appropriate and constructive use? It seems that any data feedback system alone may not be the solution to improvement of educational provision, even if accompanied by increased

investment in this sector. It may be that a better quality of feedback system would make a difference here.

Feedback of data to schools is an important component required for improvement in education provision. Unfortunately, data disseminated to schools are often not used, or used inappropriately by the schools (Schildkamp & Teddlie, 2008; Wohlstetter, et al., 2008). This inappropriate use or non-use highlights the need for in-depth knowledge of the characteristics, conditions and dynamics that may enhance appropriate use of feedback based on data from school monitoring systems. It is this phenomenon of use of feedback that was investigated in this thesis and the knowledge thus gained may lead to improved educational delivery by the schools involved in the project. This knowledge may also provide a steppingstone to enrich approaches to performance[8] augmentation in education.

However, the question of how to enhance performance in schools is complicated by contextual factors. Some authors such as Fullan and Dalin (in Visscher, 2002, p. 52), state that because the situational factors vary greatly from school to school, no general system to feed data back to schools can lead to much benefit. As Fullan (in Visscher, 2002, p. 52) explains: "[t]here is no silver bullet", there is no solution which will fit all contexts.

The South African system is not only different from other educational contexts, but there is also great diversity within it. The learners come from multiple contexts with highly variable levels of skills and knowledge when they enter formal schooling. Once in formal schooling, they also develop at different rates. The severe shortage in qualified teachers further complicates the educational improvement efforts (Department of Education, 2006d). The system is also still grappling with the aftermath of apartheid and so is forced to deal with a myriad of social problems, with significant government spending in this sector still mostly aimed at rectifying the lack of infrastructure. While the

[8] Performance in this case is not only seen as the outcome of learner performance (as it is traditionally defined), but also of how the school fulfils its role as educational provider through evidence-based practice.

vast educational investment since 1994 has led to improvement of the educational infrastructure and enrolment, unfortunately, learner performance has not shown a concomitant improvement. This re-emphasises the need to combine infrastructural improvements with evidence-based practices[9]. The complexity of the context (see Chapter 3) makes developing an appropriate feedback system to facilitate use of data in schools a challenging task. Notwithstanding this complexity, the country cannot afford to be paralysed by all these obstacles. It is essential to develop a trustworthy, viable feedback approach to facilitate the use of learner performance monitoring, while remaining cognisant of contextual factors. The feedback system must be accompanied by design principles that allow the approach to be adapted to various contexts (The issue of analytical generalisability is explored in depth in Chapter 4). The stimulation of the appropriate use of learner performance data is an important building block in improving educational delivery and addressing inequalities.

To summarise, learner performance in South Africa is poor, whether measured internationally or systemically. The phenomenon that high levels of educational investment do not result in concomitant improved learner performance is observed not only internationally (Cassassus, 2001; Hayward & Hedge, 2005), but also in South Africa (World Bank, 2008; Department of Education, 2006a). Learner performance monitoring data for evidence-based practise are required to address the problem (Brinko, 1993; Hattie, 2005; Coe, 2002). The data can, however only have a positive impact if fed back to schools. Not all approaches to providing feedback, however, lead to improvement of educational delivery (Kluger & DeNisi, 1996). Schools often do not know how to use data appropriately, may not understand it or be unwilling to incorporate it in their decision-making process (Hattie, 2005; Schildkamp & Teddlie, 2008; Wohlstetter, et al., 2008). Contextual factors

[9] Controversy surrounds the issue of evidence-based practice in education with some associating it with a particular type of research associated with the Cochran Foundation. In this thesis, the term relates to adapting planning and educational practices in schools based on feedback of data from learner performance monitoring. The definition means pedagogical practices are interrogated based on the evidence of monitoring as feedback, and thus become more amendable to change based on this evidence. This is also sometimes referred to as data-driven decision-making, which relies on systematic analysis of data and application of the analysis (Schildkamp & Kuiper, 2009).

also play a large role in determining if a feedback system will succeed (Fullan & Dalin in Visscher, 2002, p. 52). This thesis attempts to address these issues by developing a contextually appropriate feedback system for learner performance data. The focus in development throughout was on facilitating understanding and use of the feedback by schools. Part of addressing this issue was exploring why data are not used or are used inappropriately. The thesis not only resulted in the optimising of a feedback system for the South African context, but also in design guidelines to support other designers in developing effective learner performance feedback systems.

Aims of the Study

The aim of this study was to identify and understand the characteristics of an effective feedback system and the utilisation thereof in order to design and optimise a feedback system that facilitates the use of learner performance data in South Africa within the school environment. The focus was on the management and classroom levels, specifically on principals, HoDs and teachers. This study of the feedback investigated use of both the processes of monitoring and feedback and the data generated by these processes.

The aim of the study was therefore two-fold:

1. To enhance, optimise and contextualise a learner performance feedback system
2. To identify design principles and characteristics of an effective learner performance feedback system

The first aim was achieved by using a design research approach to optimise an existing feedback system (SAMP). This was accomplished through gradual and successive cycles of design, implementation and evaluation of the feedback system prototypes. The focus of the design process shifted from contextualising the system, through establishing conditions for use of the feedback, to establishing finally how schools use the feedback.

The second aim was achieved by consulting existing literature surrounding the documented characteristics of an effective feedback system. This was

followed by an investigation into how to achieve the optimal conditions to allow for use of the feedback system in the South African context. The data were generated through evaluations of the feedback system prototypes. The processes schools employed to transform the feedback into planning and action in schools were also investigated, to adapt the feedback system to be appropriate for the current users. The process culminated in the development of design guidelines for an effective learner performance feedback system based on the literature review and data generated through the design research process.

The primary focus of the study is to enhance the feedback system and identify the associated design guidelines to facilitate use. This investigation includes both the identification of the optimum conditions for use of feedback that a monitoring facilitator should endeavour to establish, as well as the study of the processes in schools for transforming the feedback into action. The development of a feedback system and relevant design principles may cascade into change of facilitator, principal and teacher behaviour surrounding the use of feedback. The effects of this study may contribute to the improvement of teaching and learning in the wider South African context, as well as providing guidelines to feedback designers in different contexts.

1.4 Research Questions

The general research questions are presented in this section, to be operationalised in Chapter 3 and explored further based on the literature review and conceptual framework. The overall research question is:

- What are the characteristics of an effective feedback system and the use thereof in designing an optimum feedback system to facilitate appropriate use of learner performance monitoring in primary schools in South Africa?

Characteristics refer to the elements that should be present in such a feedback system, as well as how these should interact with each other in the context for which it was designed. The various characteristics form a gestalt where the interaction of the characteristics forming the whole is more than the

sum of the parts. An optimum feedback system, relevant to the context and needs of the users, must:

- achieve consistency between the various elements of the feedback system
- be practical for the users
- be effective in informing planning and practise for improving learner performance

A feedback system that facilitates use produces not only understandable and accessible data for schools, but also encourages and stimulates the use of the data for evidence-based practice. The learner performance monitoring in this study refers to data generated through the SAMP assessments on Grade 1 learners' performances in English Additional Language, Handwriting, Early Phonics, Early Reading and Early Mathematics. The definition of use incorporates not just instrumental or direct use of the data by schools, but also conceptual use or development of skills and approaches for working with data, planning and problem-solving that may result from participation in the feedback system.

In order to address the overall research question, it is necessary to examine a number of specific sub-questions. In this study there are six questions needed to answer the main question.

1. How can an existing learner performance monitoring system be appropriately adapted, contextualised and translated to the South African context?

The optimisation of the feedback system necessarily involves a parallel process of improving the implementation of the school-based monitoring system, SAMP. Although this improvement of SAMP is not the central focus of this study, it is necessary to establish that SAMP has been appropriately adapted and contextualised to the country to ensure it generates valid and reliable contextually appropriate data. The credibility of the assessment also contributes to the schools' perceptions of the data and likelihood that the

feedback will be used in the schools. Research sub-question 1 is addressed in Chapter 2.

2. What are the characteristics documented in literature of an optimal feedback system for use in school-based monitoring?

Existing knowledge reflected in the literature had to be incorporated in the design of the optimal feedback system for SAMP. Sub-question 2 helped to identify the essential characteristics of a feedback system for school-based monitoring as indicated by the literature. Further information about what is necessary for optimal use of the feedback was gained through the close study of other school-based monitoring systems that have led to effective use of data in the school environment. This study included an examination of the current South African educational context (see Chapter 3) in which the system functions. This information provided the basis for the global or overall design for the SAMP feedback system. Research sub-question 2 was the focus of Chapter 3.

3. What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?

Once the documented characteristics of an optimal feedback system were identified in the literature, it was necessary to establish how to optimise the various components of the SAMP feedback system. The primary focus was on investigating how the optimal conditions for use of the feedback system could be established. Research sub-question 3 was the focus of Chapters 5-8.

4. How do schools use feedback?

The purpose of the feedback system was not only to provide feedback to schools but also to facilitate use of the feedback for decision-making, planning and action in the schools. In order to understand fully how the feedback system could facilitate the use of the feedback in schools, it was important to study how feedback was being used in the schools. Research sub-question 4 was addressed in Chapters 7-8.

5. How effective is the feedback system in enhancing classroom practices, management and planning activities?

This question brings the study to the issue of praxis, or directed change and to what extent the feedback was transformed into action and, consequently, improvement in the school environment. In order for a feedback system to be effective, it must have a measurable impact on the quality of education in terms of classroom, management and planning activities in schools. In order to measure causal links in improvement in learner performance, an experimental design is required that has the potential to be an extensive study. Such a study is beyond the scope of this thesis, but evidence of changes in classroom, management and planning was documented through questionnaires, observations, interviews and monitoring data. This evaluation therefore includes both perceived efficacy as informed by input of principals, HoDs, teachers and academic consultants as well as some evidence of actual efficacy. Research sub-question 5 was addressed in Chapters7 and 8.

6. Which design guidelines for the development of an effective feedback intervention for school-based monitoring can be identified?

Finally, it is essential that there is a wider impact than this specific study and context. The design principles are required to identify the characteristics of an effective feedback system to address the main research question and contribute to the body of knowledge on school performance feedback systems. In order to provide a basis for transferability of this research to different contexts, design principles are identified from the design research process for use in other contexts. These principles may support other researchers in their own development of monitoring feedback systems that enhance evidence-based practice. Research sub-question 6 was addressed in Chapters 3 and 5-8.

1.5 Research Methodology

This research aims not only to generate knowledge by describing the characteristics of an effective feedback system and developing design guidelines, but also to design and develop a well functioning feedback system. The main research question lends itself to a design research

approach[10] that aims to align research and utility (De Villiers, 2005; Van den Akker, 1999). The design research process is iterative and follows a cyclical pathway of development (Nieveen, 1997; Richey, Klein, & Nelson, 1996; Thijs, 1999). As De Villiers (2005) explains, design research is a cyclical iterative analysis of design, development and implementation, combined with formative evaluation to understand the issues of the application domain. In this study, for example, each cycle of design research consists of the design and introduction of a version or prototype of the feedback system. This in turn is formatively evaluated, leading to a further cycle of development with a new prototype. Developing various prototypes is seen as generating "successive approximation of the ideals" (Van den Akker, 1999, p. 2). The prototypes are aimed at providing solutions in a real life context.

The prototypes in this study are gradual, successive approximations of the ideal feedback system for SAMP that encompass both the knowledge and the application domain. In this study four new prototypes were developed. Firstly, the feedback system design process was concerned with how to present and disseminate the data, as well as how to provide support to facilitate the understanding and interpretation of data by schools (Prototype I and Prototype II). Secondly, it was concerned with how to stimulate the appropriate use of the data for educational planning and application in schools (Prototype III). The final prototype (Prototype IV) along with the accompanying design principles based on the development process were the products of the study. Therefore, both the application and the research components were addressed.

The design research approach for this study incorporates various combinations of qualitative and quantitative methodologies during each evaluation cycle of the prototypes. As such, this study is located in the pragmatist paradigm, foregrounding issues of utility above those of method and propagating the use of the most appropriate tools to investigate a

[10] Design research was previously known as development research, as the research aims to design or develop a real world solution to a problem. Van den Akker (1999) in his earlier works referred to development research. Currently authors refer to design research (Plomp, 2009), while some prefer the term 'design-based research' (Joseph, 2004). In this thesis, the term 'design research' is used for consistency.

phenomenon (Onwuegbuzie & Johnson, 2004, 2006; Onwuegbuzie & Leech, 2005; Tashakkori & Teddlie, 1998). Pragmatism was a highly appropriate paradigm for this study as the study specifically focuses on the use of feedback (The choice of paradigm is discussed in full in Chapter 4). This design research study applied mixed methods, combining quantitative and qualitative approaches across all phases of the research, from conceptualisation to inference. Onwuegbuzie and Johnson (2006) describe this type of design as fully integrated mixed research and note that such an approach is attractive due to the multiple points of integration and complementarity it involves (The overall research methodology is discussed in Chapter 4).

1.6 Presentation Style

I have multiple roles in this research, including both monitoring and feedback facilitator and evaluator during the design research process. I explore these roles and implications for this research in depth in Section 4.3.1 and explain the role of self-reflexivity in managing these roles and the tensions.

Throughout the thesis some research diary and memo entries will be used for reflexivity and to make my disposition and process clear to the reader. Other than in these sections, I will refer to myself in the third person as 'the researcher' - a personal choice to facilitate the writing and reading process and provide consistency. This approach is appropriate since the research is a design research process and whilst I acknowledge my signature will have an influence on the research, it is not an emphasised component of design research.

1.7 Structure of this Thesis

The remaining chapters in this thesis are introduced below along with a short description of the content of each chapter.

- **Chapter 2: The South African Monitoring system for Primary schools**

This Chapter addresses research sub-question 1. The SAMP instruments provide the data and processes for which this study is developing and

improving the feedback prototypes. It is essential to have a firm grasp of the monitoring system itself. The aim of the chapter is to provide information about the development and South African contextualisation of the instruments. It concludes with a description of the quality of the monitoring data that are generated by the SAMP instruments.

- **Chapter 3: Contextualisation, literature review and conceptual framework**

This chapter addresses research sub-question 2. The feedback prototypes for this research were designed for the South African schooling system and policy context. The chapter starts by describing the country's education system and concludes with design criteria for the design of the prototypes in that context. This chapter also provides a rich description of the South African context to afford the readers an opportunity to transfer or adapt this research to their own contexts. Four prominent international school information systems are described. This provides a steppingstone to the development of the feedback prototypes. The chapter examines literature from the fields of monitoring utilisation, feedback, school improvement and effectiveness to develop the conceptual framework that guided the study and interpretation of the findings.

- **Chapter 4: Overall Research Design**

This chapter starts with a discussion of the ontological, epistemological and methodological basis of this study. The design research approach applied in this study is explored. The chapter concludes with a discussion of the methodological norms of the study and ethical considerations. Only an overview of the research design and the choice of the research approach are discussed in this chapter. Detailed descriptions of the design and methods for each cycle are provided in Chapters 5-8 as each successive cycle is discussed.

- **Chapter 5: Preliminary Phase: Problem identification, needs and context analysis**

This chapter focuses on research sub-question 3. The aim is to examine the problem in context and establish the criteria for an optimal feedback system. The chapter explores an exemplary case study of a proven and effective feedback system, namely the asTTle (assessment tools for teaching and learning) in New Zealand (NZ). The NZ context was compared to the South African context and design criteria for the SAMP feedback system were generated. A detailed discussion of the research procedures for this cycle is combined with the presentation of the data and design principles.

- **Chapter 6: Prototyping Phase: Establishing conditions for use (Cycle 1-2)**

This chapter documents and discusses the design, procedures, data and results for the first two design cycles. Sub-questions 3-4 are the focus of this chapter. The cycles aimed to establish conditions for the use of the feedback and launched the investigation into how schools used the data. Data collection included expert evaluations, the Delphi technique and questionnaires. Two design prototypes were developed and evaluated. Specific sampling, analysis and research procedures for the two cycles are explored. The chapter explores the evolution of the design principles from one feedback prototype to the next.

- **Chapter 7: Prototyping Phase: Transforming conditions for use (Cycle 3)**

This chapter documents and discusses the design, procedures, data and results for the third design cycle. The focus of this cycle is on sub-questions 4 and 5, although elements of sub-question 3 are also expanded. The cycle aims to establish how conditions of use are transformed into classroom, planning and management practices in schools. Data collection included questionnaires, observations, structured reflective journals and interviews. Specific sampling, analysis and research procedures for the cycle are explored.

- **Chapter 8: Assessment Phase (Cycle 4)**

This chapter documents the semi-summative evaluation of the fourth and final prototype for this thesis. The design, procedures, data and results for the cycle are discussed. The focus of this cycle is sub-questions 6, developing the design guidelines for the study. Sub-questions 2-4 are also explored for the final prototype. Data for the semi-summative evaluation were collected through expert evaluation reports and questionnaires for teachers and school management.

- **Chapter 9: Conclusions and recommendations**

The final chapter presents a summation of the findings of the research and the conclusions to be drawn from this study. In particular, it explores the implications of the design principles for the SAMP feedback system as well as the applicability of these findings to other contexts. A set of conclusions and recommendations is presented, along with a discussion of the possible effects of these findings on policy, practice and research. The chapter therefore addresses the overall research question. The limitations of this study are explored along with recommendations for further research.

CHAPTER TWO

The South African Monitoring system for Primary Schools

> An implicit assumption of almost all school effectiveness research is not only that pupil progress is good but that the schools that encourage the most rapid progress are the best. This largely untested assumption seems most likely to hit problems for very young children starting school. (Tymms, et al., 2000, p. 105)

The first research sub-question is addressed in this chapter: How can an existing learner performance monitoring system be appropriately adapted, contextualised and translated to the South African context? The chapter provides information on how the original PIPS monitoring instruments were developed in the UK, then translated, and contextualised for South Africa to become SAMP. The adaptation, translation and contextualisation of the Vocabulary subtest of the PIPS instruments are used to illustrate this process. The data quality of SAMP in the South African context is explored. The validity and reliability of data produced for the sample of schools through the SAMP instruments are established to show that the data could be used with confidence to examine use of the feedback system in the South African context.

2.1 The PIPS Instrument

The CEM was established in 1983, based at the University of Durham, it is the largest university-based educational research unit in the UK (Curriculum Evaluation and Manangement Centre, 2007). When the National Curriculum was introduced in the UK in 1988, it led to a formidable public accountability system for English schools. The accountability system ranges from the publishing of league tables to regular inspections by the Office for Standards in Education (OFSTED). The OFSTED findings are also published and made available to the public on an OFSTED website (Tymms & Albone, 2002). The CEM developed out of a drive to improve the education system from within i.e.

changing accountability practises to be more improvement-orientated through the existing structure.

The CEM is rooted in the concepts of 'distributed research', ownership and participation by practitioners to solve educational problems (Fitz-Gibbon, 1996; Tymms & Coe, 2003). The CEM's professional monitoring systems are for use by practitioners and not meant as official accountability systems that may lead to blaming of practitioners (Tymms & Coe, 2003).

The CEM has developed a number of monitoring systems that are used by about 7,000 schools in the UK and assess the progress made by over a million learners every year (Tymms & Coe, 2003). The CEM suite of monitoring systems (see Table 2.1) caters for learners from 3-19 years (Curriculum Evaluation and Manangement Centre, 2007; Tymms & Coe, 2003). This suite of value-added assessments was developed in conjunction with teachers and Local Education Authorities (LEAs) staff in the UK. The assessments combine objective assessments and teacher ratings to provide valuable information about each learner (Tymms & Albone, 2002).

For the South African adaptation, two points in schooling were identified as crucial areas where monitoring was necessary, namely entries into primary and secondary school (Howie, 2002). South African schools often receive learners with highly diverse backgrounds and levels of skills from wide feeder areas when entering primary and secondary education (Scherman, et al., 2006). At the time, there were also no South African monitoring systems in place that focused specifically on these transitional points. This meant adapting the CEM's assessments for the beginning of primary school (Performance Indicators in Primary Schools or PIPS) and the beginning of secondary school (Middle Years Information System or MidYIS).

Table 2.1: CEM information systems

ASSESSMENT	AGES	FOCUS
A Level Information System (ALIS)	16-19 years	Vocational and academic Value-added measures up to Advanced level examinations.
Year 11 Information System (YELLIS)	14-16 years	The baseline for GCSE grades, comprising mathematics, vocabulary and perceptual reasoning. Value-added scores can be calculated and additional information captured.
Middle Years Information System (MidYIS)	11-16 years	This is a curriculum free assessment examining mathematics, vocabulary perceptual reasoning, proofreading and perceptual speed and accuracy.
Performance Indicators in Primary Schools (PIPS)	5-12 years Infant & Junior	Provides assessment for all primary schools, used in England and Scotland. Examines developed abilities in vocabulary and non-verbal ability as well as outcome measures in terms of mathematics and reading in order to provide concurrent value-added measures.
Performance Indicators in Primary Schools (PIPS)	4-5 years On-entry baseline and follow-up	Assesses variables shown to be good predictors of later success in schooling. Value-added is measured against Foundation Phase profile at end of reception in England.
Interactive Computer Assessment System (InCAS)	5-10 years	A personalised diagnostic assessment as part of the monitoring systems
Assessment Profile on Entry for Children and Toddlers (ASPECTS)	3-4 years	Prevent age-related feedback compiled from information from home and preschool as well as assessments of physical personal and social development as well as language and mathematics.

(Curriculum Evaluation and Manangement Centre, 2007; Tymms & Coe, 2003)

This PIPS instrument adapted for South Africa provided the basis for the development and evaluation of the feedback system prototypes in this thesis. The CEA decided to adopt PIPS, adapting it for the country's context and implementing it by means of funding from the NRF (Howie, 2002). The PIPS assessment fulfilled the CEA's criteria for an assessment measure in that it provided an indication of a child's readiness for academic learning as scores on the test administered at the start of schooling and correlated well with subsequent academic achievement (Tymms & Coe, 2003). PIPS was also administered twice a year to provide a measure of progress and the

assessment provided information on a child's profile of performance in a number of domains that could be used to identify particular learning difficulties or strengths.

PIPS was developed in 1994 for the purpose of providing a baseline by which to assess progress in reading, phonics and mathematics for learners entering primary school (Tymms, et al., 2004). The monitoring system aimed to improve education through feedback to the schools (Tymms, 1999). The baseline assessment was combined with a follow-up assessment to determine the value added[11] by schools between the two assessments.

The full PIPS instrument consists of 14 subtests, 13 of which are combined into three different scales: Early Phonics, Early Reading and Early Mathematics. The Handwriting subtest is separate and constitutes a scale score on its own. The scales are generated as follows:

1. The ***Early Phonics*** scale is derived from two subtests, the Repeating Words and Rhyming Words subtests. These subtests focus on phonic awareness as an important basis for the development of reading ability.
2. The ***Early Reading*** scale is derived from seven subtests: Vocabulary, Ideas about Reading, Letters, Words, Story and Sentences. This scale focuses on the prerequisite skills for reading development.
3. The ***Early Mathematics*** scale comprises seven subtests, namely Sizes, Counting, Sums A, Numbers and Mathematics. These subtests aim to establish the learners' abilities in early mathematics.

The baseline assessment consists of only 11 subtests, which are repeated in the follow-up assessment, during which a further three more advanced subtests are also administered. The results from the follow-up and baseline assessment are compared to determine value-added scores between the two assessments.

[11] Value-added measures quantify the value added by a school to learning for a learner, taking in consideration various factors. There are different ways to calculate value-added scores, these are discussed in depth in Section 2.2.1

PIPS was originally developed as a paper-based test, but in the UK is usually employed in its computer-delivered format (Merrell & Tymms, 2005). The PIPS assessment was administered individually and usually took 20 minutes per learner. Tymms and Wylde (2003) report that schools generally find the administration time manageable.

Following the widespread success of PIPS in the UK, it was adapted with minimal changes for countries such as Australia, New Zealand and Scotland. The Dutch and British Sign Language versions were developed to maintain the original intentions and characteristics of the items (Merrell & Tymms, 2005). Germany, Lesotho, Thailand, France and Hong Kong have also adapted PIPS (Tymms, et al., 2004). The CEA is responsible for the South African adaptation of PIPS known as SAMP.

With PIPS being implemented in several countries, the opportunity for international comparison of school entry levels of learners was created (Merrell & Tymms, 2005). The use of the data for international comparative purposes however reiterated the importance of appropriate adaptation that maintained the difficulty levels and intentions of items across applications in various countries. As the South African context differs widely from that in the UK, the unique learning context of South Africa was expected to influence how children perform on the CEM instruments. Therefore, it was necessary to adapt aspects of the monitoring system.

2.2 The South African Birth of SAMP

A number of effective learner performance monitoring systems exist (see Chapter 3 for a discussion of four such systems). The CEM instruments were chosen for adaptation as the CEM expressed interest in working with the CEA and in fostering monitoring and evaluation skills in South Africa. The CEM was also willing to provide their instruments to be adapted and implemented in the country without charge and to provide support where necessary during development. The CEM suite of instruments has also enjoyed widespread sustained success and has been adaptable to various international contexts, whilst maintaining its integrity.

The CEA and the CEM started to collaborate in 2003. The CEA decided to adapt PIPS and MidYIS for the South African context and implement them at the beginning of primary and secondary school. Funding was provided by the NRF.

The project aims to provide quality data to schools. Performance data are aggregated on school, class and learner levels allowing for monitoring of learner performance and improvement of educational provision by schools. The system enables schools to monitor their own performance outside of a formalised accountability framework. These data may also be used to inform policy development. The initial research questions to be addressed by the project were as follows:

1. How feasible is a monitoring system using value-added measures in South Africa?
2. How valid and reliable are existing value-added measures for South African education?
3. What additional assessments will need to be developed to enhance monitoring systems?
4. What is an effective means of providing feedback to schools using value-added measures?
5. What assistance will the educators need in using the results from the assessments?
6. How will the effectiveness of the monitoring system be evaluated?

This thesis extends the original project's research question four. The focus here is not only on the most effective way to provide feedback, but also on how to provide support to understand the feedback and to transform it into evidence-based action in the school environment.

The primary school project (PIPS) was brought to South Africa for a number of specific purposes. The first was to provide an indication of a child's readiness for academic learning. Scores on the test administered at the start of schooling correlate well with subsequent academic achievement (Tymms &

Coe, 2003). Secondly, the same test is administered again at the end of the first year at school to provide a measure of progress made. Thirdly, baseline scores can also be used to estimate progress in further years. Fourthly, the test provides information on a child's profile of performance in a number of domains that can be used diagnostically to identify particular learning difficulties or strengths. Fourthly, the data from the instrument can be used to adapt the classroom teaching practises to suit the needs of the group and individual learners.

Using a value-added approach to measure learner performance is highly appropriate for the South African context as it considers prior performance. The approach is more equitable and provides for a more constructive feedback than data that ignores learners' baseline skills, especially in a country where great learner diversity has historically been employed as the basis for discrimination. The importance of the value-added approach in monitoring learner performance is discussed in the following section.

2.2.1 The Importance of Value-Added Measures

As the term implies, value-added measures assess the value that a school adds to learner achievement while considering the influence of intake factors such as background, prior achievement, aptitude and abilities (Scheerens, et al., 2003). Value-added measures are generally seen as a more equitable and balanced approach to monitoring of learner and school performance when compared to measures which ignore these factors (Rowe, Turner, & Lane, 2002; Saunders, 2001). Value-added measures engage a variety of statistical methods to adjust gross output indicators to incorporate intake factors (Scheerens, et al., 2003). Bosker and Witzier (in Rowe, et al., 2002, p. 172) differentiate between three different kinds of value-added measures:

1. **Unpredicted achievement:** adjusted for family background factors and student ability.
2. **Learning gain:** adjusted for initial achievement level.
3. **Net progress:** adjusted for family background factors, ability and initial achievement.

Currently the SAMP project uses the learning gain measures for schools in the sample. The monitoring system allows for easy comparison of baseline and follow-up performance results that is necessary for a learning gains approach.

The importance of value-added measures can be illustrated by examining the factors that contribute to learner performance. As Saunders (2001) explains, raw scores reveal more about student background than about school performance. In Figure 2.1, Hattie (2005, p. 13) illustrates the role players and their relative contributions to variance in learner performance.

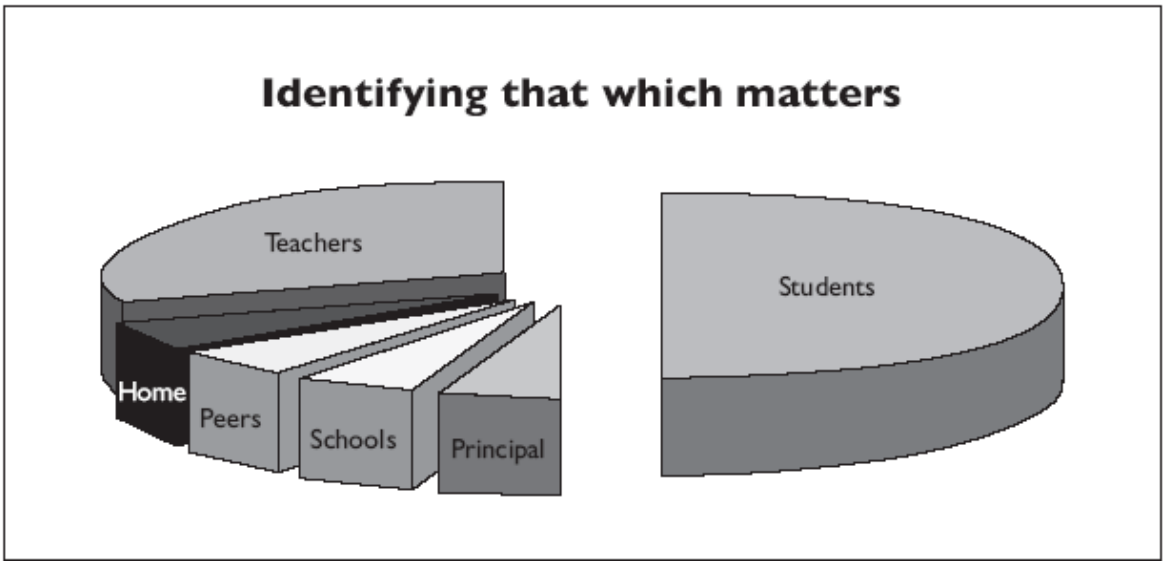


Figure 2.1: Contributions to learners' achievement variance
(Hattie, 2005, p. 13)

As the pie chart indicates, factors inherent in the learner (e.g. intelligence, maturity, socialisation, motivation and aptitude) are responsible for 50% of the variance in learner performance. Systems such as the league table published in the UK neglect intake factors inherent in the learner, which could misrepresent schools. Omitting learner intake characteristics creates the impression that the intake learner profiles of all the schools are similar or irrelevant and that any ranking is thus purely based on the quality of the school. This type of accountability system may have multiple adverse consequences:

- Political and media 'bashing' of schools and teachers
- Test-dominant curricula
- Schools selecting stronger pupils for intake to boost 'league table' rankings
- Parents taking drastic measures to enrol their children in high ranking schools

(Rowe, Turner, & Lane, 2002, p. 166)

Schools in the new democratic South Africa have also experienced media persecution. This could be seen with the listing in the press by the DoE of the worst schools in each of the nine provinces (Jansen, 2001), a practise that was only ceased under the guidance of Minister Naledi Pandor. Some countries have taken a very firm stance to prevent public comparison of schools. These countries include Ireland that introduced legislation in 1998 to prevent school-by-school comparison (Looney, 2006).

In South Africa, the need for value-added measures - as opposed to performance outcome measures alone - is accentuated by the great diversity amongst learners. Factors and conditions that influence learner performance range from issues of language and socio-economic circumstances to an educational system that is still trying to align schools under one education department (schools were categorised under various education departments during the apartheid era). Any monitoring system that remains blind to these factors will not generate quality data for feedback to schools in the South African context. The next section examines the adaptation of the PIPS assessment to show how it was contextualised and translated to the South African context to become SAMP.

2.2.2 Contextualisation and Adaptation of SAMP

The adaptation and contextualisation discussed here are illustrated through examples. In this case, the Vocabulary subtest items are used to illustrate how SAMP was adapted and contextualised. The complete, extensive process is detailed in Archer, Scherman, Coe and Howie (2010).

The first step in contextualising PIPS was to determine the aspects of PIPS that presented difficulty in the South African context and needed to be re-examined (see Section 2.2.2.1). The second step was generating suggestions for adapting the identified problematic items (see Section 2.2.2.2). Thirdly, the functioning of the newly adapted instrument had to be assessed (see Section 2.2.2.3). The vocabulary subtest is used throughout the discussion as the illustrative subtest. The vocabulary subtest was selected as it was subjected to multiple graphical and language changes during the adaptation.

2.2.2.1 The PIPS Instrument in South Africa Prior to 2006

In 2003, the UK English language version of the PIPS instrument was translated and adapted for use in three South African languages[12]: Sepedi, Afrikaans and English. These three languages represent the three dominant LOLT in the Tshwane region, the region in which the CEA is located (H. Julies, DoE, personal communication, July 16, 2010). The original UK-based PIPS instrument was translated into Sepedi and Afrikaans by registered translators, then corroborated through a process of back-translations. The translated instructions and items were recorded for the computer-based assessment by voice artists and the voice clips were sent to the CEM to be incorporated into the code of the computer-based PIPS. By 2005, the PIPS computer-based assessment was available in all three languages with learners guided through the assessment on a one-on-one basis by a fieldworker who captured answers directly on the computer as the learner provided them.

When this researcher was appointed to coordinate the project in 2006, 416 Grade 1 learners were participating in the project. This included learners from two Afrikaans, two Sepedi and three English Language of Learning and Teaching (LOLT) schools in the Tshwane area. A team of seven trained fieldworkers would travel to each school with rented laptop computers and the software to assess the Grade 1 learners. At the time, sustainability, growth

[12] All learners are assessed in the LOLT of the school and not their home languages as the assessment aims to establish readiness and ability to perform in their current academic setting. This means that many of the learners are not first language learners. The most diverse learner population in terms of home languages is found in the English LOLT schools.

potential and the cost of the project were of concern. A review of the progress of the project and planning for it further development was conducted.

The review of the project status (for full reports see Archer, 2006a, 2006b) examined the handover and process documentation from the previous project team (Eiselen, 2004; 2005b; Küstner & Eiselen, 2005) and a report by an educational psychologist (Barry, 2005) on the cultural fairness of the instrument. It soon emerged that administering the assessment on computers contributed disproportionately to the project expenses and both slowed and complicated adapting and contextualising the instruments. Theoretically, the computer-based assessment offered a number of benefits over paper-based assessment:

1. **Ease of capturing of data:** Data are directly captured on the computer during administration.
2. **Ease of administration:** In its computer-based format, the PIPS assessment automatically implements termination rules. The fieldworker is guided through all the administration procedures in a systematic fashion and the fieldworker is forced to capture the data before the next item is revealed.
3. **Interactive nature of the presentation:** As a computer-based assessment, the learners are confronted with and interactive environment. Murphy and Davidshofer (1994, p. 191) state, "[t]here is evidence that people view computerized tests more positively than equivalent paper-and pencil tests. This often contributes to the learner's ability to maintain concentration and interest in the assessment".
4. **Facilitation of international comparability:** PIPS is administered as a computer-based assessment in most of the participating countries. Computer administration thus facilitates comparison of the South African data with that of other countries. (Archer, 2006a)

These perceived benefits of computer–based administration of PIPS were not realised in the field in South Africa. The CEA was obliged to reconcile the perceived benefits of computer-based assessment with the realities of the country:

1. **Ease of capturing data**: Several difficulties with capturing the data on the laptop computers in the field were discovered and reported by Küstner and Eiselen (2005). These difficulties included incorrect entries that were difficult to correct due to an absence of hardcopies of the assessments.

2. **Ease of administration:** Many of the fieldworkers did not have the required level of proficiency to utilise a computer. Extensive training in use of the computer-based assessment was required.

3. **Faulty termination rule application:** It was found that the computer programme would sometimes terminate pre-maturely irrespective of the learner's responses. (This can be seen in the Letter subtest data where the test terminated after eight items irrespective of learner responses). It was also difficult to check for such errors, as the CEA did not have direct access to the programming code for the assessments.

4. **Incorrect insertion of sound clips**: As the PIPS assessment remains the property of CEM, the translated sound clips for the assessment in Afrikaans and Sepedi had to be sent to CEM to be inserted into the programme. This resulted in some of the clips being placed incorrectly.

5. **Cost of administration:** Most of the schools involved in this project did not have computer laboratories. Therefore, laptop computers had to be rented for the fieldwork. The cost of renting laptop computers for the fieldwork represented a major part of the expenditure in the project.

6. **Security:** Travelling with valuable equipment such as laptops presents a serious security risk in South Africa. This negatively affected the safety of the fieldworkers.

7. **Administrative burden:** The process of booking, renting, collecting and returning the laptop computers, as well as having to upload the necessary software repeatedly for fieldwork was a large administrative burden to the CEA team.

8. **Administration time:** The administration time of twenty minutes projected per child (Tymms & Wylde, 2003, ¶ 16) was greatly increased in the PIPSSA project, as laptops had to be set up and fieldworkers were often not as computer literate as the teachers in the UK are.

9. **Sustainability:** In order to achieve true sustainability for this project it would be necessary to empower teachers to administer this test and relay the data to the CEA. In order to achieve this, it is essential to ensure that the necessary infrastructure is in-place. Currently there are vast discrepancies in the availability of computer facilities for schools in South Africa. The South African Department of Education Draft White Paper on e-education of August 2003 indicates that in 2002 only 26.5% of schools had access to computers for teaching and learning (Department of Education, 2003b). Thus, the computer-based assessment would prevent certain schools from administering the assessment themselves.

10. **Adaptation for cultural fairness:** One of the CEA's objectives with the PIPPSA project was to adapt PIPS to be culturally fair for the South African context. In order to do this certain items had to be amended. Items need to be translated and various other changes were required. As programming can only be done at the CEM this entailed a cumbersome process of negotiating changes and sending changes to be programmed in the UK and then be sent back.
(Archer, 2006a)

Based on these findings, the decision was made to change from computer-based to paper-based administration. A new name, SAMP, was adopted to signify this shift. As part of the transition, the equivalence of two types of administration had to be established. The equivalence was investigated through a small sample of 96 learners. All learners were assessed twice, using the paper-based and then the computer-based formats. Half of the learners were assessed using the computer-based format first and the other half with the paper-based first to control for any learning effect.

No significant difference was found in the performance of learners overall or on any of the scales, other than the Early Reading Scale (See Table 2.2). Particular note was made of the subtests in which there was a significant difference between the two modes of administration for further investigation (Vocabulary, Ideas about Reading, Ideas about Maths and Sums A). All these subtest were heavily reliant on graphic elements and enjoyed particular attention during the adaptation process discussed below. Although the sample was small, it provided sufficient information to determine that the two modes of assessment were fairly equivalent for the selected sample. Greater equivalence would be achieved through the adaptation process. These results re-affirmed the feasibility of moving forward with the development and adaptation of the SAMP project in paper-based format.

Table 2.2: Difference between paper-based and computer-based modes of delivery of the PIPSSSA assessment

SUBTEST/SCALE	PAPER-BASED	COMPUTER-BASED	DIFFERENCE	SIGNIFICANT AT 0.1 LEVEL
Vocabulary	33.5	30.5	3.0	*
Ideas about Reading	29.0	41.0	-12.0	*
Repeating Words	66.9	66.3	0.5	
Rhyming Words	28.9	31.2	-2.3	
Letters	16.7	15.5	1.2	
Words	8.7	12.0	-3.3	
Ideas about Maths	73.0	79.8	-6.8	*
Counting	78.0	81.3	-3.3	
Sums A	54.2	48.1	6.1	*
Numbers	32.1	32.9	-0.8	
Early Phonics Scale	47.9	48.7	-0.9	
Early Reading Scale	22.0	24.7	-2.8	*
Early Mathematics Scale	59.3	60.5	-1.2	
Total Scale	43.0	44.7	-1.6	

2.2.2.2 Adaptation of PIPS into SAMP during 2006

Establishing the validity and reliability of the instrument was key to determining how well it functioned in South Africa. When adapting an instrument, the core validity issue is determining what adaptations and accommodations preserve the meaningfulness of the scores (Fuchs, Fuchs, Eaton, Hamlett, & Karns, 2000). Validity is achieved through the removal of irrelevant construct variance, created by the difference in culture, context, language, and social practices. Therefore, an important aspect of establishing the 'validity argument' (Kane, 2006) is to demonstrate that test scores measure the same thing across all groups for whom the instrument is intended.

The core validity issue in adapting an assessment for the South African context is therefore determining which adaptations and accommodations would preserve the meaningfulness of the scores (Fuchs, et al., 2000). Validity addresses the question as to what extent the interpretation of results is appropriate as well as meaningful (Gronlund, 1998). Validity is a unitary concept that is based on various forms of evidence, with construct-related validity being the central concept. Ultimately, validity is concerned with the consequences of using the assessment (Gronlund, 1998; Killen, 2003; Linn & Gronlund, 2000).

One strategy for identifying bias in an assessment is to look for differential item functioning (DIF) (Smith, 2004). If the relative difficulty of an item differs significantly across various groups, it indicates that scores that include that item are not measuring a uni-dimensional construct. This means that performance on the item is being influenced by some characteristic of that group other than the underlying construct being assessed (Smith, 2004).

Generally, reliability refers to the consistency of scores obtained by the same individuals when they are requested to complete the assessment on different occasions (Anastasi & Urbina, 1997). Reliability not only indicates how much confidence can be placed in a particular score, but also how constant the scores will be over different administrations (Owen & Taljaard, 1996). Reliability for the assessment was established by investigating internal consistency, which is seen as a pre-requisite for construct validity. High inter-item correlations are expected among items that measure the same construct (Kline, 1993).

Reliability of Subtests and Scales

The Cronbach's alpha values for the 2005 computer-based assessment (see Table 2.3) indicate that some aspects needed to be re-examined. The acceptable level for a reliability figure is determined in part by the envisaged use of the data from the assessment. If the data are to be used for decision-making for a group, the figure can be lower than when it is used for decision-making concerning individuals. In the same way, a lower reliability figure can be tolerated if the data are not used in isolation, but in conjunction with other data (Frisbee, 1988).

Table 2.3: Reliability coefficients for the computer-based 2005 PIPSSA subtests overall and according to language of assessment

SUB-TEST / SCALE	OVERALL (n = 417)	AFRIKAANS (n=62)	ENGLISH (n=211)	SEPEDI (n=144)
Vocabulary	.85	.92	.85	.63#
Ideas about Reading	.87	.88	.88	.81
Repeating Words	.84	.88	.77	.85
Rhyming Words	.86	.84	.83	.77
Letters[13]	-	-	-	-
Words	.93	.93	.89	.86
Ideas about Maths	.83	.95	.75	.87
Counting	.85	.90	.80	.89
Sums A	.88	.87	.88	.84
Numbers	.91	.92	.90	.90
Early Phonics Scale	.85	.86	.80	.79
Early Reading Scale	.92	.92	.91	.81
Early Mathematics Scale	.93	.95	.91	.91
Total Scale	.96	.95	.95	.91

- items to be investigated further as indicated by Cronbach's alpha values

For the PIPPSA assessment, reliability values of above 0.8 were aimed for. It should be noted that creating reliable assessments for very young children is notoriously difficult (Archer, et al., 2010). Over and above indicating the stability of measures over time, a high reliability figure would also strengthen the inferences made about the content-related validity of the assessment (Suen, 1990). The reliability figures for the 2005 data were encouraging, however the low reliability figure of 0.63 for Sepedi learners on the Vocabulary subtest was of concern and highlighted the need to investigate the subtest further[14]. It was necessary to determine if the low figure was related to the construction of the subtest overall, the graphical presentations used, or, for instance specific phrasing of items in the subtest. As validity is related to whether the assessment measures what is intended to, the aim of the subtest is pivotal. The vocabulary subtest aims to evaluate the receptive vocabulary of learners and consists of 23 items. Learners are asked to point out objects in three different pictures, graded according to difficulty of the visual stimuli and items:

[13] No results for the Letters subtest in 2005 are available as the test terminated prematurely after only eight items.
[14] All other subtests for all three language groups were investigated further through the means described below, the Vocabulary items are used to illustrate the process here.

1. Kitchen scene (easy items) e.g. “Can you point to some carrots?”
2. Outdoor scene (moderate items) e.g. “Can you point to a windmill?”
3. Toy shop (more advanced items – examines learners’ exposure to literature) e.g. “Can you point to a yacht?”

A termination rule is applied. The rule requires that the subtest be discontinued when a candidate supplies three consecutive incorrect answers.

The report from Barry (2005) provided some indication that the computer-based PIPS was likely to disadvantage learners who had not been exposed to cartoons, animations or three-dimensional overlays. This may have been the source of the discrepancy in reliability figures for the Sepedi learners, as they were mostly from relatively poor socio-economic areas. The individual item statistics needed to be examined to determine if the graphic presentation alone explained the low reliability for Sepedi learners, or if the individual items also needed to be revised. The item facility and discrimination values for the subtests were studied more closely through classical test theory.

Item Facility and Discrimination Values

The item facility (also referred to as ‘item difficulty’ or ‘difficulty values’) and item discrimination values were used as indicators of items that needed closer examination. Item discrimination indicates the ability of an item to differentiate between high and low achievers. Item discrimination values of 0.25 or higher were aimed for when examining the item-total correlation values. Facility values show the percentage of learners who correctly answered the items. These values are presented separately for the first two pictures used in the Vocabulary subtest (see Table 2.4). As the termination rule was applied in the vocabulary subtest, most candidates were not presented with the most difficult items in picture 3. For the purposes of calculating item facilities, these missing items are treated as incorrect. Only the first 17 items are discussed here, for illustrative purposes.

Table 2.4: Difficulty and discrimination values for items overall and across languages

ITEM	ENGLISH			AFRIKAANS			SEPEDI		
	n	Facility	Discr	n	Facility	Discr	n	Facility	Discr
1- carrots#	211	92.9	0.23#	62	98.4	0.21#	144	82.6	0.39
2 - the knife#	211	79.6	0.39	62	98.4	0.21#	144	93.8	0.44
3 - a fork#	211	90.0	0.31	62	98.4	0.21#	144	97.2	0.44
4 - a cupboard#	207	72.0	0.35	61	90.3	0.18#	141	63.2	0.33
5 - some cherries#	198	59.2	0.36	61	61.3	-0.19#	141	59.7	0.25
6 - a pan	197	63.0	0.46	61	93.5	0.28	141	91.7	0.31
7 - a bowl#	183	48.8	0.32	59	58.1	-0.14#	139	75.7	0.33
8 - the butterfly	159	12.8	0.57	57	17.7	0.89	132	2.1	0.33
9 - the kite	140	11.8	0.58	54	17.7	0.78	127	4.9	0.29
10 - the castle	112	7.1	0.67	44	19.4	0.85	110	0	0
11 - the wasp	39	4.3	0.62	14	19.4	0.85	10	-	-
12 - the pigeon	26	5.7	0.61	13	14.5	0.82	5	-	-
13 - the windmill	17	3.3	0.66	12	14.5	0.91	3	-	-
14 - the turtle	14	5.2	0.70	12	17.7	0.89	3	-	-
15 - the violin	13	2.8	0.64	10	12.9	0.81	3	-	-
16 - the padlock	11	2.8	0.62	11	12.9	0.85	1	-	-
17 - the toadstool	11	0.9	0.28	9	9.7	-	1	-	-

- items to be investigated further as indicated by discrimination values

For very easy items (high facility value) such as item 1, the discrimination value is lower than the identified 0.25. This is to be expected as both low and high achieving learners typically answered correctly. The item could therefore only differentiate between very low performing learners and other learners. These items were maintained, as it is necessary to make allowances for very low performance learners to achieve a sense of accomplishment and to allow learners to become comfortable with the assessment.

From the item discrimination values, it seems that two items had a negative item discrimination value for the Afrikaans learners:

- Item 5 - Can you show me some cherries?

The Afrikaans translation of cherries is *kersies,* which is a homonym for birthday cake candles and cherries, both of which appear in the picture. This meant that learners often indicated the birthday candles instead of the more difficult item of cherries. Fieldworkers indicated that the same problem

occurred with many Sepedi learners as words in Sepedi are often borrowed from Afrikaans.

- Item 7 – Can you show me a bowl?

The approved translation of bowl as *papbakkie* for Item 7 of the Afrikaans assessment seemed to be problematic and an alternative, simpler translation was suggested to address the negative discrimination value.

The fact that most of the items were significantly more difficult for the Sepedi learners than English or Afrikaans learners is of concern. This may be due to the way items have been translated or graphically represented, or it may be that whilst translations are accurate, the translations are less frequently used or are more advanced words that decrease the item facility. If the graphic representations are found alien or distracting by the Sepedi learner, they may well act as confounding variables in measuring receptive vocabulary. Alternatively, the particular sample of Sepedi learners tested may have a poor vocabulary. It was necessary to explore this phenomenon further, through techniques such as Rasch analysis.

Rasch Analysis

The Rasch modelling locates both the difficulty of items and the ability of persons on a single latent trait continuum. The probability that a person of ability, β, will correctly answer an item of difficulty, δ, is entirely determined by the difference β – δ. This means that the relative difficulty of two items on the continuum is independent of the abilities of the sample of persons who have attempted them (Baker, 2001). This is a particular strength in the South African context, where the aim is to examine Differential Item Functioning (DIF) across groups whose average scores are quite different. The aim here was to establish if the *relative difficulty* of the items were similar across the three language groups. The difference in *average performance* of the three groups becomes irrelevant as only the relative difficulties of the items to each other for each group is examined. The assumption is that similar relative difficulty values should be obtained across the languages to show equivalence of the instrument across the languages.

All 23 items from the three pictures of the Vocabulary subtest were included in the Rasch analysis of the Vocabulary subtests. Only items from the first two pictures, items 1-17, are discussed here to illustrate the process. Rasch analysis copes well with the missing data, this is beneficial in examining assessments where a termination rule is applied.

A person separation reliability of 0.67 was achieved in the assessment, indicating that the scale discriminates well between persons. The items also created a well-defined variable or single underlying construct (as indicated by the item separation reliability of 0.98). The OUTFIT mean square for both persons and items were slightly more than one, indicating underfit (1.04 and 1.57 respectively). Conversely, the INFIT mean square for both persons and items were below one, indicating overfit, or that the responses are too predictable (0.83 and 0.80 respectively).

Upon inspection of the items, it was found that the point-measure correlations were all positive and above 0.3 for every item included in the analysis. However, several items are identified by misfit statistics (namely Item 4 *'Can you point to a cupboard?'* Item 5 *'Can you point to some cherries?'* and Item 7 *'Can you point to a bowl?'*. These three items, although falling within the criteria of 0.5-1.5 for productive items with regard to the INFIT mean square, did not fall within the prescribed range for the OUTFIT mean square indicating that outliers are present in the data.

Figure 2.2 represents the DIF for the three language groups, where the Y-axis represents difficulty and the X-axis the items included in the analysis. Similar ability levels can be observed for the three language groups for most of the items in the vocabulary subtest. However, differences can be identified. The vocabulary items at the beginning of the assessment were very easy for Afrikaans learners. Possibly, the kitchen as represented in the picture is similar to the kitchen in these learners' own homes. Item 10 *'Can you point to the castle?'* was very difficult for Sepedi learners to identify in comparison to English and Afrikaans learners, although these learners also found this item challenging. Item 12 *'Can you point to a pigeon?'* was easier for the Afrikaans learners than the English and Sepedi learners. These items link to exposure

to literature and stimulation in the South African context, whereas in the UK, castles and pigeons are more common. Furthermore, Item 17 *'Can you point to a toadstool?'* was by far more difficult for English learners than for the other two language groups.

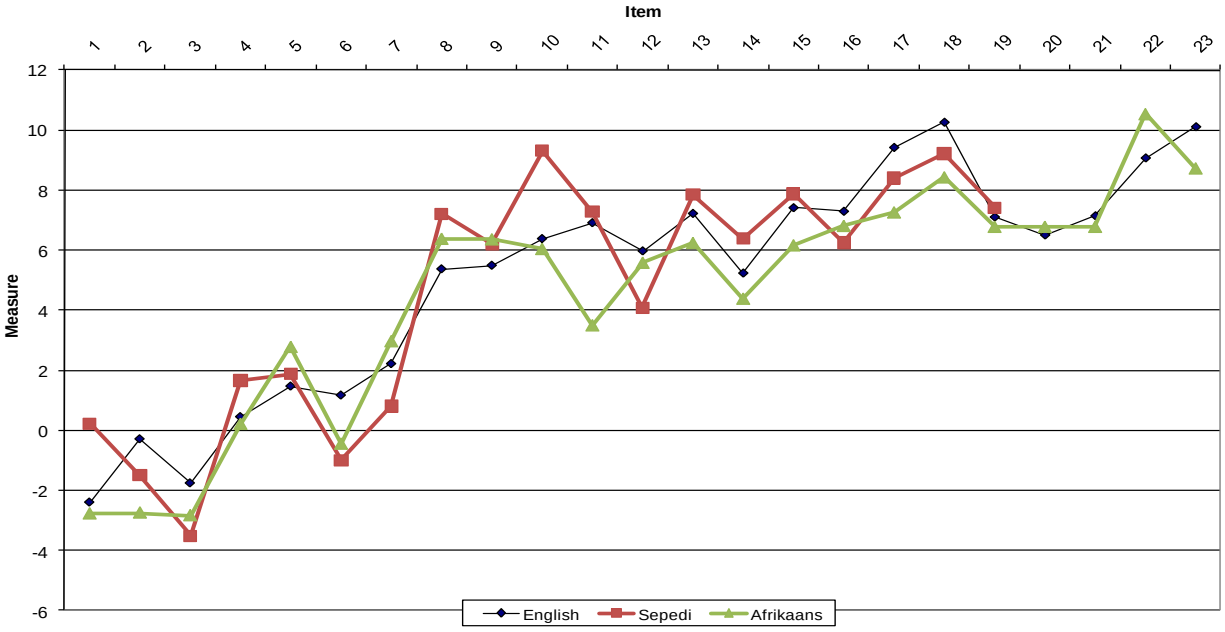


Figure 2.2: Differential item functioning for the Vocabulary subtest for the different languages groups
(Archer, et al., 2010, p. 83)

The DIF analysis indicated that some items functioned differently for the three language groups. These items were examined further to establish if the differences were due to lack of exposure to the stimulus or the translation and graphical challenges experienced throughout the adaptation process.

The statistical examination of the items from the Vocabulary subtest was supplemented by teacher evaluation of the face validity and cultural appropriateness of the items.

Teacher Evaluations

Six Grade 1 teachers (two from each language group) were asked to evaluate the vocabulary subtest by noting if the items and their graphical representations were fair in terms of exposure and culture. They were also asked to rate the difficulty of each item. The results, according to language group for the first two pictures, are seen in Table 2.5.

Table 2.5: Difficulty and values indicated by teachers and assessment of fairness of items

ITEMS	DIFFICULTY			FAIRNESS			
	EASY	AVERAGE	DIFFICULT	CULTURE		EXPOSURE	
				YES	NO	YES	NO
1 - carrots	✪✪ ◎◎ ❖❖	.	.	✪✪ ◎◎ ❖❖	.	✪✪ ◎◎ ❖❖	.
2 - the knife	✪✪ ◎◎ ❖❖	.	.	✪✪ ◎◎ ❖❖	.	◎◎❖❖	.
3 - a fork	✪✪ ◎◎ ❖❖	.	.	✪✪ ◎◎ ❖❖	.	◎◎❖❖	.
4 - a cupboard	✪✪ ◎◎ ❖❖	.	.	◎◎❖❖	.	✪✪ ◎◎ ❖❖	.
5 - some cherries#	◎ ❖	✪✪ ◎◎	❖	◎◎ ❖	❖	✪✪ ◎◎❖	❖
6 - a pan	✪✪ ◎◎ ❖	❖	.	✪✪ ◎◎ ❖❖	.	◎◎ ❖❖	.
7 - a bowl	✪✪ ◎◎ ❖	❖	.	✪✪ ◎◎ ❖❖	.	◎◎ ❖❖	.
8 - the butterfly	✪✪ ◎◎ ❖❖	.	.	✪✪ ◎◎ ❖❖	.	◎◎ ❖	.
9 - the kite	✪✪ ◎◎ ❖❖	.	.	✪✪ ◎◎ ❖❖	.	◎◎ ❖	.
10 - the castle	✪✪ ◎◎ ❖❖	.	.	✪✪ ◎◎ ❖❖	.	◎◎ ❖	.
11 - the wasp	✪✪ ◎◎ ❖❖	.	.	✪✪ ◎◎ ❖❖	.	◎◎ ❖	.
12 - the pigeon#	◎ ❖	✪✪ ◎	❖	✪✪ ◎◎ ❖❖	.	◎◎ ❖	❖
13 - the windmill	✪✪ ◎ ❖	◎	❖	✪✪ ◎◎ ❖❖	.	◎◎ ❖❖	.
14 - the turtle#	✪✪ ◎	◎ ❖	❖	✪✪ ◎	◎❖❖	◎	◎❖❖
15 - the violin	◎◎ ❖❖	.	✪✪	✪✪◎❖❖	.	◎❖❖	.
16 - the padlock#	✪✪ ❖	◎◎	❖	✪✪ ◎ ❖	◎	❖	◎◎❖
17 - the toadstool#	✪✪ ◎❖	◎	❖	✪✪ ◎◎ ❖	.	◎◎ ❖	❖

◎ = English, ✪ = Sepedi and ❖ = Afrikaans

The teacher evaluations of the vocabulary subtest raised issues about the fairness of several items. The cherries item (Item 5) was questioned by one of the Afrikaans teachers, reaffirming the difficulty highlighted by the item analysis. The validity of the graphical representations of Items 12, 14, 16 and 17, (the pigeon, the turtle, the padlock and the toadstool) were also questioned. In some cases of the teacher evaluations, there seemed to have been a tendency to confuse the difficulty of the item with fairness[15]. These cases were discussed with the teacher evaluators to clarify the issue. There was some agreement between the DIF analyses and teacher judgements. The teacher evaluations were supplemented with an expert panel review.

Expert Evaluation Panel

The expert evaluation panel consisted of two research psychologists, two educational psychologists, three teachers, two educational researchers and

[15] The term 'fairness' as opposed to 'validity' was used in the teacher evaluation documents. The document explained that fairness indicates whether the item is appropriate and reasonable to ask, given the children's culture and exposure in their context.

two subject experts involved in teacher education at a tertiary institution. The panel was presented with the above information on the Vocabulary subtest. Particular attention was paid to items flagged by the analyses and teacher evaluations. Concerns were raised by the panel that some aspects in the vocabulary were too Eurocentric, which may have acted as extraneous distracters to South African learners. The possible confounding graphic issues included:

- **Colouring:** Some learners found the colouring alien and distracting. For instance, the colouring used in the kitchen is not commonly found in South African kitchens. This may have confused learners from a poorer socio-economic status.

- **Composition and foreground-background discrimination**: The outside country scene picture was depicted as a view through a window. This caused confusion with many South African learners, as the window did not have burglar bars. Fieldworkers noted that some learners struggled to make sense of the scene, as they did not perceive it as a view through a window. The panel suggested changes to the items, paying particular attention to items flagged in the teacher evaluation.

- **Representation:** This included the windmill depicted in picture two. The windmill depicted was a Dutch windmill. South African windmills look very different with only a metal frame as base. Learners could thus know what the word windmill means, but be unable to link the word to the European representation of the windmill. This would render the item invalid as it would no longer only assess receptive vocabulary, but also exposure to European presentations of the windmill.

Table 2.6: Changes proposed to items by expert evaluation panel

ITEM	PROPOSED CHANGES
4 - a cupboard	May be unfamiliar to learners from very rural areas or with very poor socio-economic status.
5 - some cherries	May consider replacement with a South African fruit (translation into Afrikaans leads to confusion with the candles).
9 - the kite	This is a culturally specific pastime and should possibly be replaced with a more South African item.
10 - the castle	None. This may be a European concept, but learners should have exposure to this through literature. It is also a very well known South African brand name.
11 - the wasp	The drawing of the wasp is inaccurate for the South African wasp species and should be adapted. The translation of wasp into Sepedi is very complex.
12 - the pigeon	The colouring of the pigeon may have to be changed.
13 - the windmill	The item can be maintained by changing the graphic representation of the windmill to the South African windmill.
14 - the turtle	It would be more appropriate to the South African context if this item were changed to tortoise.
15 - the violin	None. Although this item demands a certain level of educational stimulation and exposure, this is true across all language groups.
16 - the padlock	The graphic presentation of this item should be changed. A grey lock with a square shape that is more familiar in the South African context.
17 - the toadstool	Toadstools are relatively unfamiliar in South Africa. “Mushroom” would be an appropriate replacement for this item, with a concurrent change in the graphic representation

Further Examination of Translation and Graphical Elements

Based on the statistical analyses, teacher and expert panel appraisals, the identified items were explored in order to determine how to address the concerns. There was some correspondence between the conceptual as well as empirical processes, with some of the same items being highlighted (items 4, 5, 12 and 17). There were also differences, e.g. the turtle, which functioned well across the groups according to the Rasch analysis and classical test theory, but which was identified as problematic by teachers. It was thus necessary for recommendations and statistical analysis to be weighed against each other in order to decide on the most appropriate adaptations. It was also important to ensure representations were seen as fair and appropriate by teachers to engender trust in the instruments and data.

The first phase was a re-examination of translations. Although a strict protocol of translation and back-translation was followed, there were still some difficulties in translation. Whilst translations were correct, they were sometimes more complex than the original text, thus increasing the difficulties of items in the translated languages.

The translation for Sepedi was very challenging. Since group names instead of specific differentiated words are often employed in Sepedi, some of the translations from Sepedi were academically correct but not often used in the spoken language. Regional dialects of Sepedi are also prolific, leading to complications in the translation process. Careful re-evaluation of the difficulty of the translations was undertaken with a number of translators. In most cases, the translations could be rectified. In the case of the word for 'wasp', no appropriate translation with a similar difficulty value could be identified and a completely new item of a similar difficulty value had to be incorporated into the Sepedi subtest.

After consultation with the CEM, it was determined that some of the more difficult items in the South African test (such as the cherries, saxophone and microscope) were quite advanced in the UK context as well and should not be altered purely because they were more difficult. The problem experienced with the homonym of cherries for Afrikaans learners was addressed by removing the candles on the birthday cake and introducing an extra distracter in the form of a box in the kitchen picture. Alterations to the colouring in the pictures made the items more accessible to Sepedi learners. All graphical elements that had to be changed to be more appropriate for the South African context were first trialled with learners to ensure that the new representations were recognisable to the learners. The adaptation of the instruments resulted in the contextualised SAMP instrument.

2.2.2.3 SAMP 2008

The shift from a computer-based assessment to a paper-based assessment in 2006 proved highly successful. This was true not only for facilitating the adaptation, but also for reducing the operational costs. As a result, the sample of learners in the SAMP project has increased since 2006. The current SAMP

sample is drawn from public schools with Sepedi, Afrikaans or English LOLT in the Tshwane region. Practical constraints such as funding and geographical proximity limited the size and location of the sample. This population therefore represents the accessible population from which the sample was drawn (Best & Kahn, 2006). Multi-phase sampling took place (Cohen, Manion, & Morrison, 2000):

1. Schools where stratified according to LOLT.
2. Eight schools where selected randomly from each LOLT from the DoE databases. A sample of 22 schools was selected on this basis, including two dual LOLT schools.

The sample was inspected at this stage to ensure a fair geographic representation of the Tshwane area and was found to be satisfactory. Currently, some 1,600 learners participate in SAMP each year and are tracked from the baseline to follow-up assessment.

The reliability figures for the SAMP 2008 assessment were well above 0.85 and were comparable to the PIPS UK reliability figures (see Table 2.7). None of the items of the SAMP assessment in any of the LOLT had a negative discrimination value in the 2008 dataset. The inclusive and transparent adaptation process had resulted in an instrument that functions well in the South African context and produced valid and reliable monitoring data that could be used with confidence to inform educational improvement through feedback.

Table 2.7: Reliability coefficients for the three SAMP scales follow-up 2008 and PIPS scales for the UK

SCALE	SAMP 2008	PIPS SCALES UK
Early Phonics	.89	-
Early Reading	.96	.97
Early Mathematics	.95	.90
Total	.97	.98

(Curriculum Evaluation and Management Centre, 2002)

2.3 Conclusion

This chapter provides information on how the PIPS system was successfully contextualised and adapted for use in South Africa for Afrikaans, English and Sepedi. The adaptation process consisted of several aspects, which included the examination of the reliability of subtests and scales; Item facility and discrimination values; Rasch Analysis; Teacher and expert panel evaluations.

The data resulted in several items and aspects of the assessment being flagged for further examination and adaptation. The data from the different processes were weighed against each other and decisions were made for adaptation (see Table 2.8 for a summary).

The reliability values of the newly adapted SAMP instrument indicate that a high quality of data can be generated through the monitoring system. These SAMP data are used in the feedback system that is optimised by this study. Due to the rigorous contextualisation, translation and adaptation process, the data from this system can be used with confidence in Afrikaans, English and Sepedi in the Tshwane region.

Table 2.8: Summary of contextualisation and adaptation of Vocabulary subtest

ITEM	RELIABILITY			ITEM STATISTICS			RASCH ANALYSIS			TEACHER EVALUATION			EXPERT PANEL			ADAPTATION
	English	Afrikaans	Sepedi	English	Afrikaans	Sepedi	English	Afrikaans	Sepedi	English	Afrikaans	Sepedi	English	Afrikaans	Sepedi	
									Kitchen scene							*Colouring altered.*
1- carrots				×	×											✓
2 - the knife					×											✓
3 - a fork					×											✓
4 - a cupboard					×		×	×	×				×	×	×	Colouring altered and 1 door opened slightly.
5 - some cherries					×		×	×	×		×			×	×	Candles removed from cake replacement distracter added.
6 - a pan													×	×	×	✓
7 - a bowl					×		×	×	×							Translation in Afrikaans changed. Translation altered.
									Outdoor scene							*Colouring changed. Burglar bars included.*
8 - the butterfly			×													✓
9 - the kite													×	×	×	✓
10 - the castle							×									✓
11 - the wasp													×	×	×	Graphic presentation and colouring altered. Sepedi item changed to item of equivalent difficulty.
12 - the pigeon								×			×					Graphic presentation and colouring altered.
13 - the windmill													×	×	×	Graphic presentation and colouring altered.
14 - the turtle										×	××					✓
15 - the violin													×	×	×	✓
16 - the padlock										×	×		×	×	×	Graphic presentation and colouring altered.
17 - the toadstool							×				×		×	×	×	Graphic presentation and colouring altered.

CHAPTER THREE

Contextualisation, literature review and conceptual framework

> … education [should be seen] as a complex system embedded in a political, cultural and economic context… It is important to keep in mind education's systemic nature, however; these dimensions are interdependent, influencing each other in ways that are sometimes unforeseeable. (United Nations Children's Fund (UNICEF), 2000, p. 4)

The chapter provides some of the global design guidelines for the optimisation of the feedback system and helps to identify the characteristics of an optimum feedback system. Section 3.1 examines the South African educational landscape in which the study is situated. The issue of use of feedback from monitoring is discussed and the study is contextualised in terms of the South African monitoring context (Section 3.2). This is essential as this thesis employs a systems approach that recognises that the use of feedback is firmly situated within a specific context. Four international School Performance Feedback Systems (SPFSs) are also examined in Section 3.3. The chapter combines literature from the fields of monitoring utilisation, school effectiveness and improvement (Section 3.3 - 3.4) to develop the conceptual framework (Section 3.5) to guide this study and the interpretation of the findings.

3.1 The Educational Landscape in South Africa

The educational landscape in South Africa forms part of the context in which feedback is used. In the following sections, the changes in that landscape since the fall of apartheid are examined. The structure of the schooling system with specific reference to the Foundation Phase of education (where SAMP is employed) is also introduced, along with a quick overview of the Foundation Phase curriculum.

During the apartheid era, education was utilised to socialise children into their expected societal roles according to race. A total of 19 different educational departments separated by race, geography and ideology were established reinforcing the inequalities of that society through their curricula (Department of Education, 2002c). The inequality was large, with the government spending up to

nine times more on the education of each white learner than on one learner from the worst off Bantustans Education (Department of Education, 2002c).

The changes in the South African education landscape since the fall of apartheid are briefly introduced in Sections 3.1.1 through 3.1.5. The discussion emphasises specific measurable indicators that influence learner performance, such as: resource availability, challenges to school attendance, the impact of social problems, issues of diversity and educator related issues. Section 3.1.6 examines the structure of General Education and Training in South Africa, and introduces the Foundation Phase curriculum.

3.1.1 Resource Availability

As far as educational resources are concerned, learners have indicated a perception of improved conditions in schools between 2002 and 2007 (Statistics South Africa, 2004, 2008). Learners perceived the greatest changes in the decrease of school fees, the increased availability of books and improvement in facilities (Statistics South Africa, 2004, 2008). This perceived improvement in the educational infrastructure has been reflected in the data in the *Report on the school register of needs* 2000 survey (Department of Education, 2001) as the National Education Infrastructure Management System (NEIMS) reports in 2007 (Department of Education, 2007b) and 2009 (Department of Education, 2009).

The data from these reports provide specific indications of improved facilities at schools from the period 1996 to 2009, which seem to support learners' perceptions of improvements in the school environment. The data have been collated and represented in Figure 3.1.

Clear improvement is shown in facilities at schools between 1996 and 2009 in terms of availability of electricity supply, water supply, sewerage removal and telecommunications. Despite these improvements, it should be noted that these figures could be deceptive, as for example they include ventilated pit latrines as a sewerage removal system and a school is viewed as having telecommunications when it has a two-way radio or a payphone on the premises (Department of

Education, 2009). Crime and vandalism[16] have had a serious impact, with theft of computers and damage to infrastructure at schools during the same period, as can be seen in the figures for laboratories and computers for learners. In 2006 (Department of Education, 2007b) 56% of schools indicated that more than 10% of their learners were still without desks.

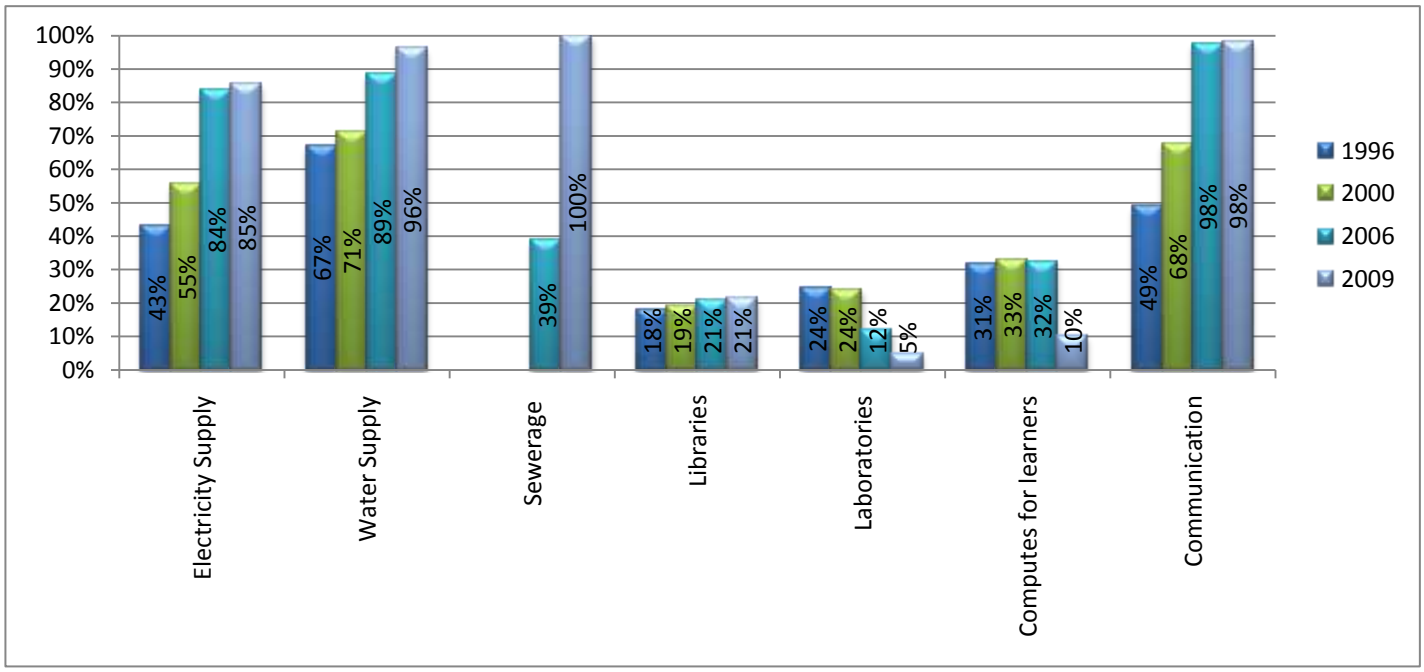


Figure 3.1: Infrastructure changes in the South African education system 1996-2006[17] (public ordinary schools)
(Department of Education, 2001, 2007b, 2009)

Large between-school variance in infrastructure exists between South African schools. Schools in South Africa are categorised into quintiles for each province based on rates of income, unemployment and illiteracy in catchment area. Quintile 1 represents the poorest schools, while quintile 5 represents the least poor schools. The quintile system is used to allocate funds differentially to schools in order to redress the large difference between schools (Van den Berg & Burger, 2002).

[16] In 2006 (Department of Education, 2007b) 32% of schools reported problems with vandalism.
[17]

Electricity Supply:	Municipal, solar, generator
Water Supply:	Municipal, borehole, rain harvesting, communal, mobile tanker
Sewerage:	Municipal flush, septic flush, Enviro Loo, VIP (ventilated pit latrine), buckets or chemicals
Libraries:	Presence of room for library, whether stocked or not
Laboratories:	Presence of stocked laboratories
Computes for learners:	Ratio of computers for teaching and learning can be one computer to over a hundred learners
Communication:	Cell phone, telephone, fax, two way radios, internet or pay phone

Although differential funding policy based on the quintile system has lead to some improvement, “the legacy of apartheid and poverty persists in terms of very varied learning contexts in the public school sector” (Consortium for Reseach on Education, Access, Transitions & Equity (CREATE), 2009, p. 4). The impact of socio-economic status and poor infrastructure on learner performance can also be clearly illustrated using the quintile system, with a large difference between the performance of quintile 5 schools and schools in the other quintiles. Performance scores (Mathematics) in quintile 4 schools are 15-30% higher than in quintile 1 schools, while performance scores in quintile 5 schools are 50-75% higher than in quintile 4 schools (Organisation for Economic Co-Operation and Development (OECD), 2008).

Therefore, although it is clear that great improvement has been made with infrastructure development since 1994, there is still great variance in school infrastructure and learners’ socio-economic status. The poor or lacking infrastructure and resources still have a negative impact on learner performance and make accessing meaningful education a challenge. The resource availability must also be taken into consideration in the design and implementation of the feedback, as one that relies too heavily on school ITC infrastructure will not be feasible or sustainable.

3.1.2 Challenges to School Attendance

South African learners face a number of challenges to attending school, the most prominent of which is finances (Statistics South Africa, 2004). This was still the case in 2007. A slight decrease in this challenge to attendance was noted from 2002 to 2007 (from 39.6% to 34.6%) (Statistics South Africa, 2008). The perception of learners that finances pose a barrier to learning also decreased after 2002 (17.8%) to 2007 (7.0%) (Statistics South Africa, 2004, 2008). This may have been influenced by the increased introduction of no-fees schools around the country, aimed at allowing for larger access to education. In 2006 there were 7,687 no-fee schools catering for 2,556,550 learners increasing to 13,856 no-fee schools with 5,001,874 learners in 2007. This latest figure represented over 40% of schools in the country (Department of Education, 2006c).

It seems that access to education in South Africa is improving, as indicated by the increase in enrolment figures particularly in the 5-7 year old age group between 2002 and 2007 (Statistics South Africa, 2008). A startling revelation as to the reasons for

non-attendance was the view amongst about 8% of learners from 2002 to 2007 that education was useless (Statistics South Africa, 2004, 2008).

3.1.3 The Impact of Social Problems

The influence of social issues on learner performance in South Africa cannot be negated. The DoE (2006d, [12]) states: “The persistence of poverty and unemployment, the debilitating effects of illness and premature death, (especially as a result of HIV and AIDS), and the threats to our environment are among the important challenges faced by the nation at the beginning of the 21st century.” HIV/AIDS presents a major barrier to school attendance and the quality of teaching and learning. This is highlighted in the DoE (2003a) document on developing HIV/AIDS policies in which it is asserted that even with resources, regular attendance and the new curriculum, the widespread illnesses and deaths in the country are negatively affecting the quality of education provided in schools. The magnitude of the problem is illustrated by the statistic in 2004 of more than a quarter of learners aged 7 to 15 years not having a parent living at home (Statistics South Africa, 2004) and more than 36% of all age groups not having a parent living at home.

3.1.4 Issues of Diversity

In general, schools in South Africa have dealt with great diversity since the fall of apartheid. Although Black African people constitute the majority of the population (Statistics South Africa, 2001), there are many White South African, Indian and Coloured learners. These population groups are by no means homogenous, with different cultures and socio-economic levels represented in each.

As far as home language is concerned, and with eleven official languages recognised in South Africa, nearly a quarter of the people indicated their home language to be IsiZulu, followed by IsiXhosa and Afrikaans (Statistics South Africa, 2001). This diversity of the learner population is a source of enrichment for learners sharing the same classroom, but may provide further challenges for educators in catering for the educational needs of all learners.

3.1.5 Educator Related Issues

There are concerns about educators in South Africa, particularly concerning the high teacher attrition rate and the quality of teacher education. In 2006, the DoE produced a document entitled: *The national policy framework for teacher education and development in South Africa: more teachers; better teachers*, in which several difficulties in providing educators were noted, including:

- poor skills profiles of educators,
- a predicted educator shortfall of 15 000 by 2008[18],
- a shortage of educators capable of teaching in African languages and
- the failure of in-service education programmes to improve education provision in South Africa.

(Department of Education, 2006d)

3.1.6 General Education and Training in South Africa, the Foundation phase

Basic schooling in South Africa from the reception year, Grade R (4-5 years) to Grade 9, falls under General Education and Training (GET). Schooling up to grade 9 is compulsory and on the completion of Grade 9, learners are awarded their GET Certificate (Department of Education, 2002c). The last three years of schooling, Grades 10-12 are not compulsory, but are required for tertiary education. The GET band is divided into three phases: Foundation, Intermediate and Senior Phase.

Each phase consists of a number of Learning Programmes aimed at integrating and attaining the learning outcomes for the specific phase. The Foundation Phase covers Grade R to Grade 3 and is the entrance into formalised education. The three Foundation Phase Learning Programmes are Literacy, Numeracy and Life Skills. The phase is aimed at establish basic skills so that learners can learn and work more independently in the later phases. The second phase is the Intermediate Phase (Grade 4-6) and the last phase is the Senior Phase (Grade 6-9). The Intermediate and Senior Phases have eight Learning Programmes: Languages, Mathematics, Life

[18] The concern about educator availability is also reflected in learners' increased concern about the lack of teachers from 2003 to 2007 from (4.3% to 6.0%) (Statistics South Africa, 2008).

Orientation, Arts and Culture, Natural Science, Economic and Management Sciences, Social Sciences and Technology. (Department of Education, n.d.) According to the Revised Curriculum Grade R-9 (Department of Education, 2002c) the developmental outcome should be an ability to reflect on and explore a variety of strategies to learn more effectively, as well as to participate as responsible citizens in the life of local, national and global communities.

The SAMP system is employed at the start and end of Grade 1 in the Foundation Phase. This provides schools with information on the intake characteristics and growth of learners in terms of phonics, reading and mathematics during this crucial phase. This covers two of the three learning areas for the Foundation Phase, which constitutes 75% of the time allocation for the phase (see Table 3.1).

Table 3.1: Time allocation for Foundation Phase

Learning Programme	Time (%)
Literacy	40%
Numeracy	35%
Life Skills	25%

(Department of Education, 2002c, p. 17)

It is within this complex educational landscape that monitoring and feedback in schools take place. In the following section, the role of monitoring and feedback in this landscape is discussed.

3.2 Monitoring and Feedback in South African Education

The monitoring culture in any country is influenced by the political environment and the country's experience of participation in previous monitoring and feedback endeavours. The focus of monitoring in South Africa is currently on performance measures, having participated in several international comparative studies such as TIMSS, PIRLS, the Second Information Technology in Education Study (SITES) and the Monitoring Learning Achievement (MLA) study (Jansen, 2001). Feedback from these studies has taken the form of reports and press releases, but the results have

been disheartening. Recently the government has made a decision to participate in fewer international comparative studies and the country did not participate in TIMSS 2007, a high profile study that usually receives much publicity and media coverage. The DoE has stated that the country was not participating in order to allow the interventions that were put in place to take full effect (Human Science Research Council, 2006). While this may be a reprieve from the extra demand of participating in another international comparative study, it did not contribute new information on learner performance in South Africa (Human Science Research Council, 2006). South Africa will however still participate in the Southern African Consortium for Monitoring Educational Quality (SACMEQ), which, in contrast to the other international studies, focuses on the Southern and Eastern African region. It may be that South Africa will compare a bit more favourably to the SACMEQ countries than those in the TIMSS study, which included countries such as New Zealand, England, USA, Norway and Australia. (Martin et al., 2000, 2004).

South Africa's monitoring culture is highly politicised, especially in terms of school monitoring and evaluation. In 2003, a resolution was made by the South African government to develop an Integrated Quality Managements System (IQMS) (Education Labour Relations Council, 2003). IQMS integrates the previous systems of Performance Appraisal, Performance Measurement and Whole School Evaluation into one process. In their report to the portfolio committee on IQMS the Parliamentary Monitoring Group describe IQMS as a paradigm shift for South Africa:

> The IQMS signals a new approach to performance evaluation in the South African education system. From an education perspective the past evaluation systems were seen as negatively focused, backward looking, judgmental, subjective, unreliable and to have a top-down orientation. The new approach, therefore, presents an opportunity for the department to turn these negatives into positives and begin to build a quality education system. (Parliamentary Monitoring Group, 2006, ¶4)

Jansen (2001) nevertheless notes that the commitment to a participatory evaluation process of the IQMS was deceptive. Educators and principals are only involved in the process of the evaluation and excluded from any decision-making. The final arbitrator in school grievances is the Minister of Education. Although IQMS purported to be focused on support and improvement of education in practice, it resulted in an old-school punitive inspectorate system. This led to what is described by Jansen

(2001) as a 'credibility crisis' for the policy and such experiences with IQMS may contribute to a sense of guardedness against monitoring and feedback from South African education practitioners.

The IQMS is the responsibility of the Offices for Standards in Education (OFSTED) (Gauteng Department of Education, n.d.) that was established in 2001. OFSTED is tasked with the monitoring and evaluation of standards of education in South Africa. One of OFSTED's specific responsibilities is to conduct systemic evaluation of learner attainment in key skills for Grade 3, 6 and 9 learners in South Africa (Gauteng Department of Education, n.d.). The systemic evaluation is aimed at allowing for benchmarking, fulfilling public accountability functions and mobilising national efforts to raise standards. Results of the systemic evaluations have however been poor, reiterating results from participation in international comparative studies (Jansen, 2001; Department of Education, 2002b, 2006a, 2006b). The adoption of OFSTED by South Africa indicated a shift in the importance attached to raising standards in education and signalled the approach that was being followed to achieve these standards.

In the recent past, the accountability system for South African education extended to include what Hattie (2005) refers to as "name, shame and blame tactics". These tactics became evident in the media exposure of underperforming schools in South Africa with headlines such as "Outrage as all matrics at school fail" (Sukhraj, 2006) and "Rasool ready to crack the whip on bad schools" (Powell & Maritz, 2006). The DoE had previously initiated a practise of naming the worst school in each province, in a sense 'blacklisting' them (Jansen, 2001). This action was based on Grade 12 marks with little consideration of the various contexts in which schools function and a blatant disregard for the legacy of apartheid that was still apparent in many of the institutions (Jansen, 2001).

During this period of external accountability pressure (2000 to 2003) there was an increase of 15.4% in the National Grade 12 pass rate (Department of Education, 2002a, 2003c, 2004) that seemed to indicate that the tactics were effective. The quality of the National Senior Certificate however came under fire. Questionable aspects included the artificial inflation of Grade 12 results by dissuading weaker students from writing the exams (Taylor, 2007), introduction of Continuous

Assessment (CASS) to contribute 25% to Grade 12 final results (Singh, 2004) and a drop in the cognitive demand of exam papers as indicated by Umalusi (Muller, 2005).

The pollution of the Grade 12 results through control of the number of learners enrolled for the exams is graphically illustrated in Figure 3.2 below. In this case, the high-stakes assessment practices lead to a negative change aimed at manipulating Grade 12 results. It does however provide an indication that monitoring and feedback can be used as a powerful lever to bring about change in the education system. The monitoring and feedback must however be applied in such a way as to encourage positive change as opposed to test manipulating behaviours. This study aims to determine how the feedback from monitoring can be used to bring about positive change in teaching and learning practices.

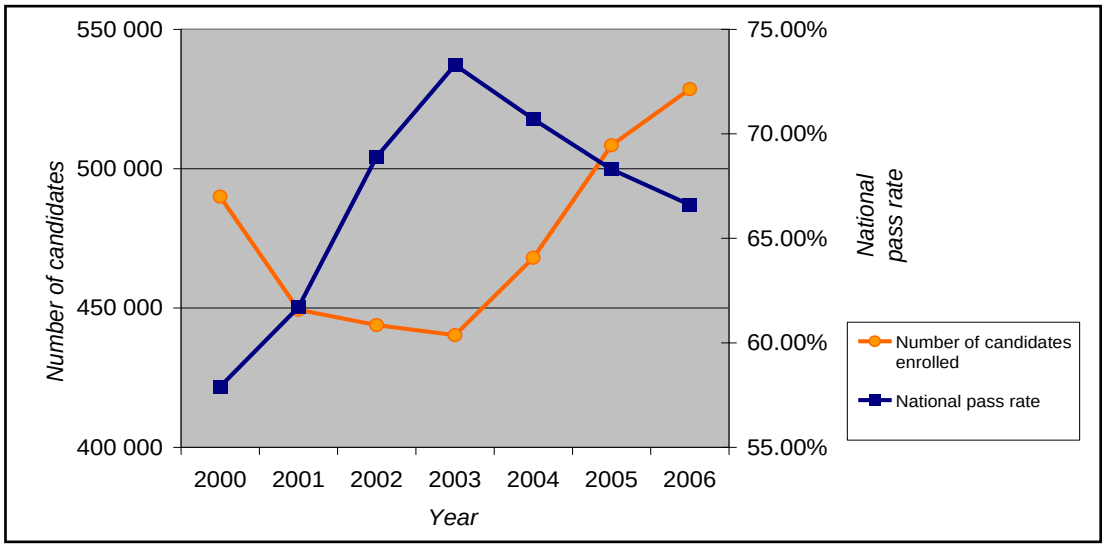


Figure 3.2: Grade 12 pass rate per year mapped against with number of candidates enrolled
Compiled from: Department of Education, 2002a, p. 181, 2003c, p. 222, 2004, p. 183, 2005a, p. 204, 2005b, p. 225; Pandor, 2006

Minister Pandor ceased these name, shame and blame tactics leading to a decrease of the annual Grade 12 pass rate and increase of Grade 12 exam enrolment for the 2003 to 2006 period (Department of Education, 2005a, 2005b; Pandor, 2006). Umalusi concluded that during this period there was an improvement in the cognitive challenge in all Grade 12 question papers (Pandor, 2006). The decrease in Grade 12 pass rate may thus be indicative of the higher cognitive demand and may indicate a levelling of the effects of CASS since 2001. With the improved cognitive demand of Grade 12 papers and decreased incentive for test pollution to avoid public shaming,

the pass rate from the latter years may well be a more valid indicator of quality of education and is arguably less likely to provide artificially inflated data.

Umalusi, the quality assurance body for Education in South Africa recently implemented changes, starting with examination of the Grade 12 papers from 2001 (Muller, 2005). This has developed into an extensive moderation process, with Umalusi now also monitoring the conduct of the Senior Certificate examination, moderation of marking and CASS. The introduction of Umalusi's Statistical Working Group has brought greater depth to the analyses of examination papers and results (Umalusi, 2007). This was further bolstered by the introduction of its Quality Assurance of Assessment and Statistical information and Research units in 2008 (Umalusi, 2009). Some of these influences on the monitoring and feedback culture as discussed are represented in Figure 3.3.

The evaluation and monitoring culture in South Africa has in part been framed by the historical and educational context of the country, which has participated in both International and regional studies of educational performance with disappointing results. Although the IQMS was introduced to replace the punitive school inspectorate system, schools have often experienced their participation in the process as merely symbolic (Jansen, 2001). The education system also went through a period where public shaming was used as an informal accountability system (Jansen, 2001), which lead to many schools participating in activities to make the results look better than they were and so to avoid the high-stakes consequences. The ‘"name, shame and blame" tactics have since been discontinued and an OFSTED established to ensure the standard of national examinations. This history has created a culture of distrust of monitoring and evaluation activities that frames how educational data are received and viewed by schools.

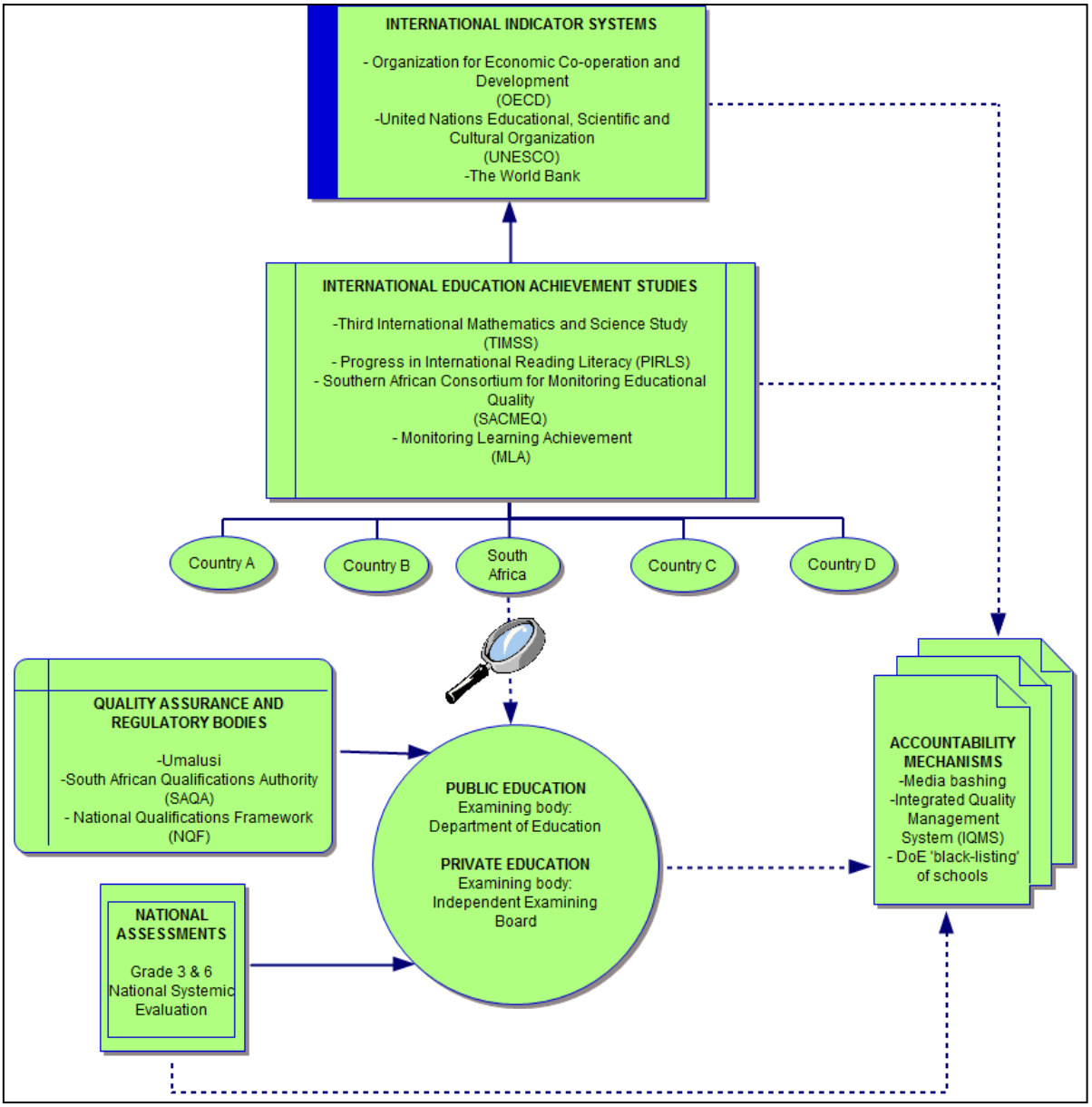


Figure 3.3: Some influences on the South African monitoring and feedback culture

3.3 Literature Review

Various systems for providing monitoring information in education for the purpose of evidence-based practise in education exist. Section 3.3.1 explores the various aims that monitoring systems may have, followed by Section 3.3.2. in which the researcher examines some of the different types of monitoring and feedback systems. As the SAMP feedback system is classified as an SPFS, some international examples of SPFS are discussed in Section 3.3.3. Literature trends on use are presented in Section 3.4 to inform the development of a framework for the

use of monitoring feedback and to identify the characteristics of an optimal learner performance monitoring feedback system. The conceptual framework for this study is presented and discussed in Section 3.5.

3.3.1 Possible Purposes of Monitoring and Feedback Systems

The most salient way to differentiate between the various systems used to generate educational data is to examine their purposes (Table 3.2). Monitoring systems and feedback systems may serve a combination of the purposes listed.

Table 3.2: Purposes and functions of information systems

PURPOSE/ FUNCTION	EXPLANATION
1. Mirroring	Description of countries may lead to identification of difference, which may help address aspects such as curriculum and achievement.
2. Benchmarking	The data may serve as a standard for comparison of achievement and progress of an education system against other countries with similar contexts
3. Monitoring	More advanced than benchmarking. Involves regular assessment of subject areas to provide trend data and to inform decision-making.
4. Understanding	This function serves to highlight the differences between educational systems. This helps inform decisions relating to issues such as the distribution and utilisation of resources and approach to teaching
5. Enlightenment	This function recognises that a general dispersal of ideas into the thinking of the community of decision makers which may not directly lead to policy. These ideas however enlighten the assumptions of policy shapers.
6. Cross-national research	International studies may indicate differences between various educational systems which lead to other studies into the effectiveness of education
7. Integration	This function has emerged from the 1980s and is related to the inclusion of previously isolated education system into the global discussion. South Africa is a prime example of such a previously isolated education system.
8. Development of less developed countries	There are four benefits of international studies: i. Development of research capacity ii. Collection of baseline data where it was not previously available iii. Establishment of a national baseline highlights what other countries are currently doing and allows for learning from other countries iv. Access to international and not merely national data gives researchers greater credence to promote education as a priority in a developing country where many issues jostle for attention of policymakers.
9. Accountability	Motivating educational expenditure, enriching discussion and reporting of the state of education as well as setting goals and performance standards.
10. Diagnosis	Identifying causes of weaknesses and strengths in the educational system or smaller educational units.
11. Decision-making	Facilitating administration and management of education, curriculum planning and classroom decision-making. For example distribution of resources, facilities, time and personnel. Allows data for planning and implementation as well as impact assessment of efforts (formative role).
12. Advancement of science	Development of theories and methods to measure outcomes on multiple dimensions
13. Administrative control	Influencing structures of decision-making in the educational system.

(Howie & Plomp, 2005; Husén & Tuijnman, 1994; Nuttal, 1994; Plomp, Howie, & McGaw, 2003; Zuzovsky, 1994/1995)

3.3.2 Types of Monitoring and Feedback Systems in Education

The feedback system optimised in this study to facilitate use is based on the SAMP monitoring system of student performance. Scheerens, et al. (2003) identifies multiple types of educational evaluation and monitoring sources that may form the basis for feedback. Monitoring and feedback systems differ not only in terms of the sources of data employed, but also in terms of the aims of these systems. The possible purposes of an educational evaluation or monitoring system are displayed in Table 3.3.

Table 3.3: Types of monitoring and educational information systems and their purposes

Monitoring Sources	Possible Purposes
Sources based on student achievement measurement	
National assessment programmes	Monitoring, Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Advancement of science, Administrative control.
International assessment programmes	Mirroring, Benchmarking, Monitoring, Understanding, Enlightenment, Cross-national research, Integration, Development of less developed countries, Accountability, Diagnosis, Decision-making, Advancement of science.
School performance reporting	Benchmarking, Monitoring, Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Administrative control.
Student monitoring system	Benchmarking, Monitoring, Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Administrative control.
Assessment-based school self evaluation	Monitoring, Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Administrative control.
Examinations	Benchmarking, Monitoring, Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Administrative control.
Education statistics and administrative data	
Systems level management information systems	Monitoring, Understanding, Enlightenment, Accountability, Decision-making, Administrative control.
School management information systems	Monitoring, Understanding, Enlightenment, Cross-national research, Accountability, Decision-making, Administrative control.
Systematic review, observation and perceptions	
International review panels	Understanding, Enlightenment, Cross-national research, Integration, Development of less developed countries, Diagnosis, Decision-making, Advancement of science.
School inspections/supervisions	Monitoring, Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Advancement of science, Administrative control.
School-self evaluations	Monitoring, Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Administrative control.
School audits	Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Administrative control.
Monitoring and evaluation as part of teaching	Monitoring, Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Administrative control.
Programme- and teacher evaluation	
Programme evaluation	Monitoring, Understanding, Enlightenment, Accountability, Diagnosis, Decision-making, Administrative control.
Teacher evaluation	Monitoring, Understanding, Enlightenment, Diagnosis, Decision-making, Administrative control.

The SAMP project is based on an SPFS developed by the CEM. This type of information system was selected as it can fulfil the functions of monitoring, enlightenment, understanding as well as diagnosis, decision-making and advancement of science and administrative control. These functions were chosen as appropriate for SAMP as the focus of this thesis is facilitating use of data for improvement of education. An SPFS can also be used for accountability purposes, but this is not the focus of SAMP.

3.3.3 School Performance Feedback Systems

Visscher and Coe (2002, p. xi) define school performance feedback systems as "...information systems external to schools that provide confidential information on the performance and functioning as a basis on school self-evaluation." SPFSs had their genesis in the wake of disillusionment with other models aimed at educational change.

SAMP was derived from an SPFS developed at the CEM. The choice of the SPFS to adapt was crucial. The strengths and weaknesses of various SPFSs are discussed here to show why this specific one was used for the SAMP monitoring system.

3.3.3.1 United Kingdom – CEM Suite

The CEM was established in 1983 and is based at the University of Durham. It is the largest educational research unit at a UK university (Curriculum Evaluation and Manangement Centre, 2007). The introduction of the National Curriculum in the UK in 1988 led to a formidable public accountability system for English schools (Tymms & Albone, 2002). The CEM developed largely out of a drive to change this educational system from within. A problem-solving instead of blaming approach is employed. The core values of the CEM include the concepts of 'distributed research', ownership and participation by practitioners to solve educational problems (Tymms & Coe, 2003). The CEM aims to establish professional monitoring systems that produce data that is actually used by practitioners (Tymms & Coe, 2003).

The CEM suite of monitoring systems caters for learners from the ages of 3 to 19 years (Curriculum Evaluation and Manangement Centre, 2007; Tymms & Coe, 2003). The PIPS assessment (that SAMP is based on) is administered using booklets. The booklets are couriered to schools, administered by the school staff and

then returned to the CEM for analysis. Feedback reports that include both charts and tables are returned to the schools (Tymms & Albone, 2002). A helpline is open for schools to gain further support.

The success of the CEM suite of instruments is seen in its widespread use, not only in UK schools, but also around the world. The PIPS on-entry baseline assessment alone is currently being used in countries such as Australia, Germany, Lesotho, Thailand, France, Hong Kong, New Zealand, Scotland, the Netherlands and South Africa (Tymms, et al., 2004). The CEM suite is a concerted move away from accountability towards using value-added measures for monitoring.

3.3.3.2 New Zealand – asTTle

In NZ, the Ministry of Education has implemented an SPFS along with the national assessment. The national assessment system takes the form of the National Education Monitoring Project (NEMP) while the SPFS is known as Assessment Tools for Teaching and Learning (asTTle) (Brown, Irving, & Keegan, 2008), both are funded by the NZ government (Ward, Hattie, & Brown, 2003). This is exceptional, as an SPFS is not usually initiated or funded by government. Developers of asTTle at the University of Auckland are vociferous opponents of accountability practices (see Hattie, 2005). Since 1989, school administration in NZ has become decentralised and schools have become highly autonomous, a context for which asTTle was designed (Crooks, 2002). Since 1999, each school has been provided with CD-ROMs that include both the assessment and data analysis tools free of charge. Schools are not forced to participate, no sampling takes place and the school performance data are kept confidential (Crooks, 2002). asTTle originally assessed mathematics and reading for English and Maori LOLT learners and rubrics for writing assessments were developed in 2001 (Glasswell, Parr, & Aikman, 2001).

The asTTle project focuses on Year 4 to Year 12 learners and consists of a number of items that have been mapped by content area experts and educators according to the NZ curriculum and curriculum content framework (Nicholls, 2003). All items are also mapped according to the Structure of Observed Learning Outcomes (SOLO) taxonomy, meaning that student performance on the various tasks can be mapped according to broad levels of current functioning (Hattie &

Brown, 2004). All items were placed in a test data bank. The educator selects specific sub-areas and levels in a subject to be assessed. A forty-minute paper-and-pencil test and a memorandum are automatically compiled by the asTTle programme.

The test results are captured after marking and the educator can then generate various types of outputs for individuals, class and school allowing for comparison across national averages according to gender, socio-economic group and an assortment of other sub-groupings. Reports are generated automatically in accordance with the preferences selected by the educator. This is accomplished through ICT calculations and displays built into the programme (Hattie., et al., 2004b). asTTle has been under continuous development since its launch in 1999, with the online version of asTTle known as e-asTTle piloted in 2008 to be made available to NZ schools in 2009 (New Zealand Ministry of Education, 2009). This will allow for automatic scoring of close-ended questions, greater flexibility in test length, as well as additional analysis and report options. Indeed, asTTle relies heavily on the technology available in schools as the test generation and analysis of data can only be conducted on personal computers. Links to a web-based learning support materials portal are provided for educators to consult, once results are available (Brown, 2007).

3.3.3.3 Netherlands - Zebo

ZEBO (*Zelf Evaluatie in het Basis Onderwijs*[19]) was developed in the autonomous Dutch school environment and when the 'Quality Law' was introduced in the Netherlands, it made schools responsible for their own quality policies to ensure improvement through self-evaluation (Hendriks, Doolard, & Bosker, 2001). No specific format for these evaluations was provided, which in some cases resulted in schools using tools and approaches of questionable technical quality (Hendriks, et al., 2001). Currently, more than 70 of these tools are available in the Netherlands (Schildkamp, Visscher, & Luyten, 2009). ZEBO was developed to address the issue of monitoring and self-evaluation by providing a quality self-evaluation instrument, the aim of the ZEBO project being to establish integrated instruments for primary school self-evaluation that creatively combined various approaches to school self-

[19] English translation: *Self-evaluation in Primary Education*

evaluation and education quality monitoring (Hendriks, et al., 2001). Schidlkamp (2007, p. 4) describes school self-evaluation as:

> a procedure involving systematic information gathering that is initiated by the school itself and intends to assess the functioning of the school and the attainment of its educational goals for purposes of supporting decision-making and learning and for fostering school improvement as a whole.

Based on school effectiveness research and school input, ZEBO uses computerised self-evaluation questionnaires for teachers, school management and pupils to generate school and classroom level process data. ZEBO only employs self-evaluation data and learner assessment data are not included. The Results are produced as classroom and school level reports that compare results for the school to those of the national average. Data are presented both graphically and textually. In the school reports, data from the school management and teachers are compared, while the school reports compare teacher and learner data. Differences in opinions by teachers and school management, or teachers and learners, are used as points of discussion by schools to achieve improvement of educational provision. (Schildkamp, et al., 2009)

3.3.3.4 America, Louisiana – School Analysis Model (SAM)

The SAM project is a product of a revised pilot programme that became part of the Louisiana accountability programme (Angelle, 2004). This is quite different from the CEM's monitoring systems and asTTle that see accountability practices as generally contradictory to improvement aims. SAM aims to introduce school process data to the accountability system in Louisiana that used to focus solely on outcome indicators, i.e. learner performance (Teddlie, Kochan, & Taylor, 2002). SAM is used for on-site school evaluations by external assessors from the Department of Education in Louisiana. Process data are collected at the classroom, grade/department, and school levels to develop context-specific, school improvement plans (Schildkamp & Teddlie, 2008).

SAM employs the ABC+ (Attitudinal/Behavioural/Cognitive Indicators, plus Context) matrix of 16 cells that encompasses quantitative, qualitative and mixed data. The matrix incorporates attitudinal, behavioural and cognitive performance indicators as well as context variables on the school, classroom, student and parent levels

(Angelle, 2004). An example of the data elements for the ABC+ model is provided in Table 3.4. The ABC+ taxonomy for school indicator system differs from other SPFSs in its emphasis on measuring ongoing processes and the importance it places on context (Teddlie, et al., 2002).

Table 3.4: Examples of data elements for the ABC+ matrix used in SAM

Level of Analysis	Attitudinal Indicators	Behavioural Indicators	Cognitive Performance Indicators	Context Variables
School (or Principal)	Interview with principal	Self-report principal's log of activities	Questionnaire assessing awareness of school change processes	Governance structure Community type Grade phase of schooling
Classroom (or Teacher)	Focus group interview with teachers	Observations in classrooms using defined protocols	Questionnaire assessing awareness of best teaching practices	Information on selection and retention of teachers
Student	Attitudinal questionnaires	Observations of a student during a whole school day	NRTs CRTs Value-added scores	SES backgrounds of students
Parent	Individual interviews with parents	Inventory of parental activities at the school	Questionnaire assessing awareness of direct parental involvement	Other cultural components of community

(Teddlie, et al., p. 85)

It is clear that these four cases have varying approaches to SPFSs, however there are certain similarities. The following common themes emerge:

Concern about high-stakes accountability practices: the CEM suite and asTTle take a firm stance against accountability, viewing such practises as contradictory to drives for improvement. The ZEBO project aims to support schools in becoming accountable through internal evaluation, while the SAM system tries to change accountability practises from within by introducing contextual and process factors.

Use of school and educator expertise in the development of the systems, for example in the development of item banks for asTTle or the prototyping and further development of systems such as ZEBO, the CEM suite and SAM.

The SPFSs aim to provide schools with ***tools to facilitate school-driven improvement practises*** and support greater school autonomy. This is accomplished through providing detailed diagnostic and management data, such as the data provided by the CEM suite and ZEBO. In the case of SAM this is also combined with recommendations, while asTTle provides links to possible courses of action through the tki[20] gateway, with its 'what next' function.

The cases reviewed also have varying degrees of ***flexibility***, be it the flexibility for schools to select only certain aspects of the SPFS, or to decide when the system is used, or even responsiveness to user input on the need for development of additional tools. In all cases a short turn around time from assessment to reporting helped to ensure the data were still relevant. In all these cases, ***ICT solutions*** were employed to facilitate short turn-around time and increase flexibility. This may be through ICT reporting engines, such as those used by the CEM and SAM to provide feedback to schools and decrease turn-around time. In the cases of ZEBO and asTTle, integrated ICT systems located at schools are used to capture data and generate several types of reports locally in real time. These types of ICT solutions also provide a greater sense of autonomy to schools.

It should however be noted that reviews of SPFSs are limited in that there is:

- ***A lack of focus on long-term outcomes of SPFSs***. The SPFSs focus on short-term change, but not on change trends over a number of years. Coe and Visscher (2002) ascribe this oversight to a concentration on timely and immediate outputs in SPFSs. This does not mean that there are no long-term outcomes for SPFSs, only that this is not currently the focus of investigations.
- Angelle (2004) expresses concern that ***little evidence is provided for the link between SPFSs and improvement in schools***. Angelle (2004)

[20] A bilingual portal-plus web community which provides quality assured educational material for New Zealand teachers, school managers, and the wider education community http://www.tki.org.nz

questions whether the provision of usable data is enough to ensure change if not accompanied by some type of plan.

SPFSs all aim to achieve change and improvement using the generated data. The PIPS system from the CEM in the UK was selected as the basis for SAMP for the following reasons:

- PIPS provides learner performance monitoring data with a high predictive value for future learner performance
- PIPS employs value-added measures that are highly appropriate in the South African context, where student intake characteristics are highly variable
- The system is not designed for accountability practices, but for improvement purposes with a focus on collaboration and ownership by schools
- The suite of instruments has enjoyed widespread sustained success
- PIPS has been adapted to various international contexts whilst maintaining its integrity
- The CEM was willing to provide their instruments to be adapted and implemented in South Africa without charge and to provide support where necessary
- The CEM indicated a desire for collaboration to develop monitoring and assessment skills in South Africa.

(Archer, 2006a; Archer, et al., 2010; Coe, 2002; Tymms & Coe, 2003)

In the next sections, the literature and conceptualisation of use of monitoring data are examined.

3.4 Use of feedback in schools

Increased school autonomy has highlighted the increased need for informed decision-making in education (Bosker, Branderhorst, & Visscher, 2007; Hendriks, et al., 2001). How data are used in schools is often still a mystery and complicates the task of feedback facilitators in providing feedback that suits a school's data-use needs.

> More recent is the attention that is given to the feedback of indicators to individual schools. More and more stakeholders become convinced of the fact that a better use of the indicators could lead to powerful opportunities for individual schools to analyse and improve their quality of education (Van Petegem & Vanhoof, 2005)

Coe (2002) explains that knowledge on monitoring data-use has mainly been derived from organisational settings, psychology and the theoretical underpinnings of education. Researchers have been grappling with feedback and data-use for decades, especially in the field of evaluation (Kirkhart, 2000). Black and Wiliam (1998) argued that performance feedback should be about particular qualities of learners and their work and how they can improve, believing that feedback should be thoughtful, reflective and focused to allow for exploration and understanding. The same principals can be applied to feedback of learner performance data to schools. Black and Wiliam (1998) advocate starting with a small group for implementation to serve as living examples to schools and educators, followed by gradual dissemination of the approach. This process should go along with a reduction in obstacles to data-use and continued research in optimising feedback.

Evidence suggests that feedback can often be harmful and may not always improve a situation (Kluger & DeNisi, 1996). Thus, when designing and implementing feedback systems one should be mindful that it might not automatically lead to improvement, but under the right conditions, feedback can have a substantial effect on the improvement of task performance (Coe, 2002). Salpeter (2004) states that

> "[t]he most important element of an effective data-driven program[me] is not the data, the analytical tools, or even the curriculum framework…it is the school culture in which the data inquiry takes place".

Schools that have problem-solving capacity, innovation capacity and attitude, extra resources allocated to data-use and tailored support and training, are more likely to show improvement due to feedback of data (Hulpia & Valcke, 2004). The school data-use culture is however not static, but can be altered through development, interaction with the feedback system and positive experiences of data-use (Schildkamp & Kuiper, 2009).

To improve the possibility of the use of data, the data should also be relevant, reliable, valid, up-to-date, anonymous or confidential, have differential effectiveness

and evoke positive reactivity (Hulpia & Valcke, 2004; Bosker, et al., 2007). Very often it is not the information itself that is of importance but rather the way in which it is conveyed (Brinko, 1993; Schildkamp, et al., 2009). Brinko (1993) advocates two-way communication and a variety of methods of feedback, including modes such as video feedback and written or face-to-face presentation to accommodate the users' preferences and learning styles. In all cases it is essential to take into consideration users' experience and development in using data. The impact of feedback depends on the interaction between the feedback message, the nature of the task performed and situational variables (Coe, 2002).

In addition to what type of information is needed and how the information is presented, another vital element relates closely to how the agent who receives the information uses it. Visscher (2002) includes the component of "use" as a central concept in the way in which he theoretically articulates SPFSs. Use lies on a continuum from direct use to use purely for informational purposes without resulting in action (Alkin & Taut, 2003; King & Pechman, 1984; Weiss, 1981). Use in the context of this research specifically refers to the process of applying the knowledge received toward a solution of a problem or the attainment of a predetermined goal (Love, 1985; Schildkamp & Kuiper, 2009). The application of the information may include direct use (instrumental use), delayed use or diffused conceptual use (Beyer, 1997; Estabrooks, 1999; Harnar & Preskill, 2007; Love, 1985).

The different types of use identified as important for this research are *instrumental*, *conceptual* and *symbolic*. Instrumental use is the concrete application of the research information in a specific and directed way (such as decision-making) (Harnar & Preskill, 2007; Love, 1985). The understanding of instrumental use can be expanded by referring to its specific purposes (Schildkamp & Kuiper, 2009, pp. 3-4):

- Genuine improvement actions
- Instructional purposes
- Supporting conversations
- Professional development
- Encouraging self-directed learning
- Policy development and planning
- Meeting accountability demands
- Legitimizing actions
- Motivating students and staff
- Personnel decisions

Conceptual use on the other hand is using the information for general enlightenment in which one's thinking may be changed but does not result in action. Symbolic use is when information is used to legitimise practice, defend a position or in a persuasive way to lobby for resources (Beyer, 1997; Estabrooks, 1999; Harnar & Preskill, 2007; Visscher, 2002). However, sometimes a decision may not be taken as a direct result of the research information but rather the research information is considered but does not drive the decision-making process. Thus, the information is used in an indirect manner (Love, 1985; King & Pechman, 1984).

Regardless of how the information is used, Weiss (1981) suggests that use should be studied in terms of what is used, who uses it, how immediate is the use and what its effect is. Utilisation is a process that entails adoption and implementation. The process itself has a number of elements (Brown & Rodger, 1999):

1) Dissemination of performance data;

2) Evaluation of the applicability of the information received;

3) Incorporating the information into practice.

The conceptual framework for use for this study is elaborated on in the next section, having been developed from the literature on evaluation utilisation, SPFSs, feedback, school effectiveness and school improvement within a systems theory framework. Monitoring is seen as a particular brand of evaluation and the literature on evaluation utilisation therefore provides invaluable insight into the phenomenon of

use of feedback from the SAMP system for the enhancement of teaching and learning.

3.5 Conceptual Framework

The conceptual framework (see Figure 3.4) developed in this section provides a structure to understand the use of the learner performance feedback system at school level. It also underpins the design of the study in order to provide a perspective for the results to be discussed. The framework identifies characteristics, conditions and processes that form part of the use of the monitoring, but does not identify how these may specifically act to facilitate or hinder the use of the data and processes. The information from the literature on how these characteristics, conditions and dynamics impact on use of feedback is shown here. These aspects are however explored further in Chapters 5-8, where the data from this research are discussed. An expanded conceptual framework based on this research is reproduced in Chapter 9. The framework employs a systems theory approach (Johnson, 1998; Patton, 1997) that acknowledges that the various components interact and impact on each other, producing complex feedback loops between the various aspects. The framework was populated through a grounded theory analysis of the literature to avoid mere eclecticism.

The concept of developing models to understand the use of feedback and educational data have a long history. As early as 1969, it was noted by Egon Guba (in Alkin & Daillak, 1979, p. 41) that "... evaluation has had little influence on educational decision-making, and evaluation information is largely ignored". Evaluators attempted to improve use of data by increasing the methodological rigour. This however, had an unexpectedly adverse effect as evaluations became more technical and expensive, rendering them less practical and understandable (Fullan, 2006; Patton, 1991). Literature on feedback, SPFSs and data-driven decision-making also reflect on how data complexity and lack of data-literacy hinder school improvement through evidence-based practise (Black & Wiliam, 1998; Hattie , 2005; Schildkamp, 2007; Schildkamp & Kuiper, 2009; Schildkamp & Teddlie, 2008; Visscher, 2002)

It soon became clear that use of data is a very complex phenomenon, with various approaches attempted to make sense of it. In the 1970's evaluators started to list the various factors and variables believed to influence use (see Alkin & Daillak, 1979). Soon it became apparent that models or frameworks, rather than lists, were required to show how these factors interacted. In response, a proliferation of evaluation utilisation models were produced, clearly depicted by Johnson's (1998) review of nearly twenty such models proposed in evaluation literature prior to 1998. Johnson (1998) concludes his review by producing a meta-model for evaluation utilisation based on the reviewed models. A similar approach has been taken in literature on SPFSs, school improvement and effectiveness and feedback research (Bosker, et al., 2007; Salpeter, 2004; Schildkamp & Kuiper, 2009; Visscher, 2002).

While the existing evaluation models provide a good basis, none of the models are appropriate for this thesis on the use of a monitoring feedback system. Many of the earlier models are linear and focus mostly on the methodology of programme evaluation, often seeing use as the final step in the programme evaluation process. The models however changed over the years, showing shifts in the utilisation literature (Shulha & Cousins, 1997) that included:

- increased realisation of the importance of context
- acknowledgment of the significance of process use
- expansion of the perception of users from the individual to the organisational level
- expansion of the role of evaluator to include those of facilitator, planner and educator/trainer

Models from SPFSs, school improvement, effectiveness and feedback in education have provided additional information on the educational variables that influence use of feedback. These model are, however, often linear and do not emphasise the multiple systems interaction that influence use of monitoring feedback in schools for evidence-based practise.

These mentioned models were employed to develop a conceptual framework for use of the feedback system associated with SAMP. In order to avoid ambiguous

eclecticism, a grounded theory method analysis of the literature was undertaken to identify the recurring themes and proposed interactions between the various factors and conditions. Some significant insights or shifts were also identified out of the literature to be incorporated.

Figure 3.4 and employs systems theory as proposed by Patton (1991; 1997) and Johnson (1998). The conceptual framework is a nested system acknowledging that any feedback is situated in the greater external education, monitoring and feedback environment in the country, as well as by the immediate, internal context, in which the feedback system and schools are functioning. The use of the feedback is in large influenced by the interaction of the feedback and monitoring culture of the schools, the characteristics of the monitoring system through which the data are generated, as well as by the characteristics of the feedback and monitoring facilitator. Combined with the characteristics and approach used in the feedback of the data, these four groups of characteristics interact to determine the dynamic or flux characteristics of the relationship between the users (e.g. schools, teachers, principals) and the monitoring and facilitator characteristics. The credibility, trust and ownership that the users feel, based on these interacting characteristics, influence how receptive the users are to using the feedback. Enlightenment (increase in understanding and knowledge) takes place either through the data provided by the feedback system, or the process of participating in the monitoring and feedback.

The information is then viewed through a political lens and tested for reasonability and feasibility before it results in the different types of use. Every time use and participation in the feedback system take place the monitoring and feedback culture and experience of the users change and influence following cycles of feedback and use. The concepts and individual interactions depicted in the conceptual framework are discussed in Section 3.5.1 and Section 3.5.2.

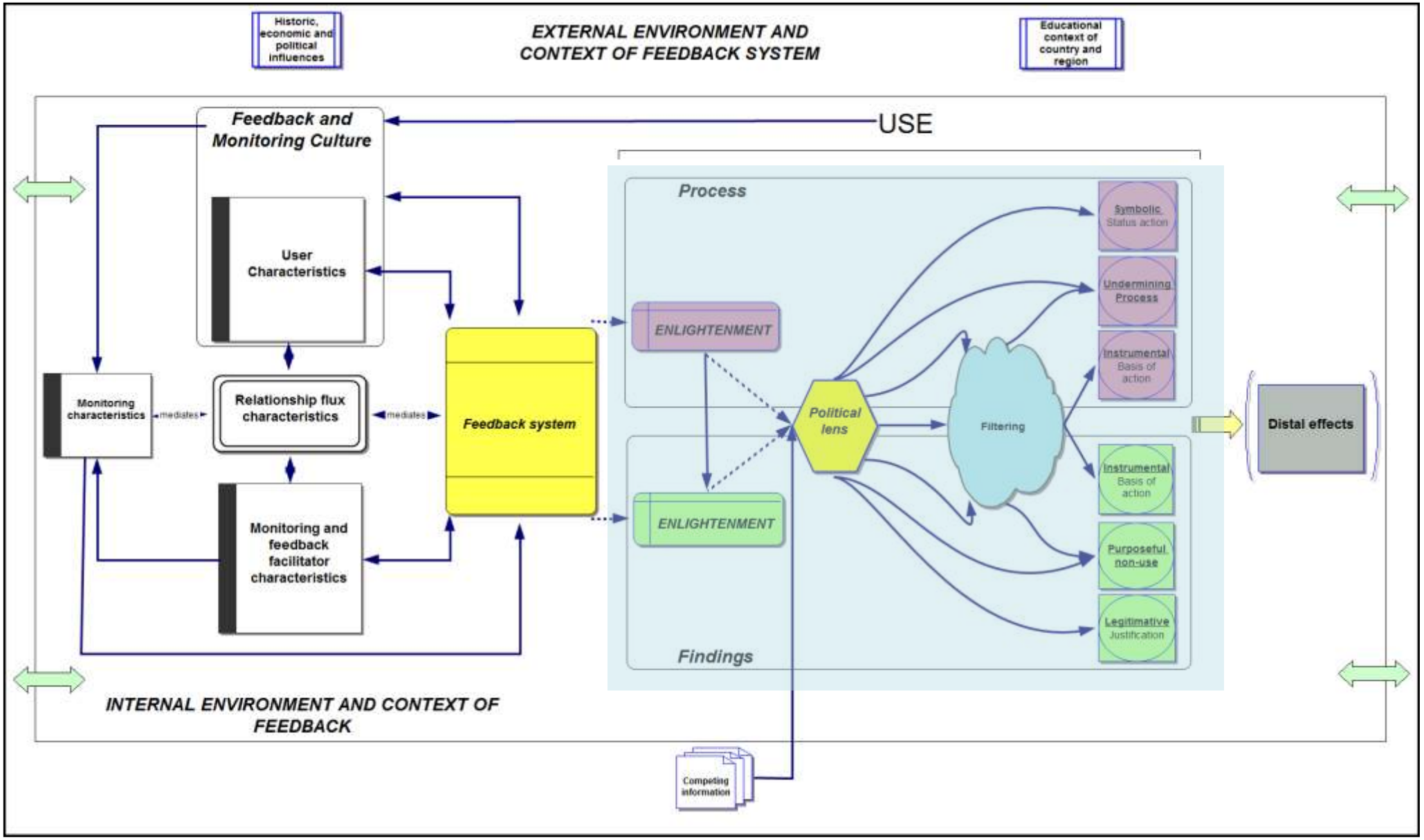


Figure 3.4: Preliminary conceptual framework for monitoring feedback use

3.5.1 External Environment and Context of the Use of the Feedback System

The use of feedback takes place within the environment and context in which school or institution are situated. These include the ***historic, economic and political influences*** in the country or region. In South Africa, this would include the legacy of a highly segregated educational system during the apartheid era and the variable school infrastructure (Howie, 2002). The ***educational context*** relates to the facilitating and inhibiting factors inherent in the educational system itself and may encompass support services available through the DoE, administrative responsibilities of educators and human resource factors. In South Africa, the high administrative demand in terms of reporting to the DoE, as well as the shortage in qualified educators falls in this category (Department of Education, 2006d).

Competing information is also available in the environment external to the school. In South Africa international and national assessments (Jansen, 2001) and the IQMS (Education Labour Relations Council, 2003) are available. There may also be research or other monitoring taking place and even everyday knowledge and media reports may provide a source of data. Most of this data emanates from the educational context, but everyday knowledge and beliefs that are rooted in the historical, economic and political context also form a part of this competing information.

Any feedback system is received in the context of these competing data sources. Either schools may integrate all the information from all or selected sources or some sources may be selectively ignored (Hattie, 2005). In cases where sets of data from the various sources are contradictory, the situation becomes even more complex. Competing information therefore influences the decisions and use during any monitoring and feedback process.

3.5.2 Internal Environment and Context of the Use of the Feedback System

Three types of characteristics constantly interact throughout the use of the monitoring and feedback process. These are characteristics inherent in the:

- Monitoring and feedback system
- Monitoring and feedback facilitator characteristics
- User characteristics

Monitoring and feedback system characteristics relate most to the relevance of the choice of methodology for the context. If users do not see the chosen methodology used for monitoring as relevant and appropriate for their needs, use will not take place.

The paradigm employed grounds the methodology and shapes how the roles of the users and monitoring facilitator are defined (Cousins & Leithwood, 1986). For example, if a constructivist paradigm is employed, as in the case of the SAMP feedback system, the feedback and monitoring are seen as a collaborative process. Users contribute their contextual expertise to the monitoring facilitator's research skills. In this paradigm it is more likely that the school users will feel a sense of ownership of the data than in a case where school users are only seen as informants.

The choice of methodology must also consider the users' needs and perceptions. Some users may be uncomfortable with certain methodologies where only vast quantities of quantitative data are generated, while other users may not place any value on qualitative data. If there is poor fit between the user and methodology, resistance to the monitoring and use of the feedback system may increase.

Users also examine the quality of the monitoring and feedback. Rigour pertains to the quality and integrity of the monitoring and feedback conducted and must be assessed using appropriate measures. For quantitative methodologies this may include reliability, validity and generalisability, while trustworthiness in terms of credibility, dependability and confirmability, as well as transferability are employed for qualitative methodologies (Guba & Lincoln, 1985) (see Chapter 4 for a full discussion). The quality of the method is judged to determine if the data are reliable and valid. Users may often not use methodological norms, but foreground issues of relevance (Schildkamp & Kuiper, 2009), in which case the design research guidelines provide a more intuitive manner of assessing the usefulness of the

feedback system, focussing on relevance, consistency, practicality and efficacy (Plomp, 2009).

Monitoring and feedback facilitator characteristics refer to the role of the facilitator, e.g. expert, planner, or educator (Alkin & Daillak, 1979; Rossi, Lipsey, & Freeman, 2004). The facilitator's skills, social competence, contextual knowledge and technical knowledge interact with personal attributes that include language spoken, culture and background, to influence the facilitator's perceived competence. These characteristics often have a greater influence on use of feedback than technical competence (Taut & Alkin, 2003).

The role of the facilitator is influenced by the paradigm in which the feedback utilisation is taking place. It is unlikely that the facilitator would conduct the monitoring and feedback in a paradigm or use a methodology that with which he or she is uncomfortable with or does not value. The same paradigmatic factors influence the researcher's view of the users' roles, which may fall anywhere on a continuum from informants to full collaborators in the monitoring and feedback (Nieveen, 2009). With SAMP and the attached feedback system, users are seen as full collaborators who influence the design and adaptation process of the monitoring and feedback system.

User characteristics refer to both the characteristics of an institution and the individual people functioning within it. This may include a school's approach to problem-solving and attitude towards change. In terms of the individual users this includes the person's personal approach to change, preferences of interaction style, personal motivations for participation, and what sort of data and monitoring are valued. A single monitoring feedback system may have to cater for several users at various levels. The personal characteristics of a user play a major role in the use of the feedback system and directly influence the quality of dialogue between the facilitator and user. What the user requires from the feedback system, as well as his or her priorities for change and commitment to change, influence the level and type of use of the feedback system. The users' locus of control influences whether feedback is transformed into action. If users distance themselves from any responsibility by transferring it to the learners or context, productive use will not take place (Schildkamp & Kuiper, 2009). Users also require a sense of agency to be able

to effect change in the school. Their characteristics interact with the monitoring culture of the organisational system (e.g. school) in which they functions.

The ***feedback and monitoring culture*** of the organisation influences how monitoring and feedback is perceived. This culture is usually moulded through prior experiences and use of monitoring and feedback and may include vicarious learning from the experiences of other organisations and individuals. The characteristics of the school leadership or principal play a crucial role in use of feedback and culture of monitoring and feedback in the school (Fullan, 2006; Visscher, 2002; Schildkamp & Kuiper, 2009; Wohlstetter, et al., 2008). The value the principal places on the data, support he or she provides, enthusiasm and encouragement for evidence-based practice is often the determining factor in how well data are received and used. Contextual factors such as the availability or perceived availability of resources may influence the perception of feedback and monitoring. The principal again plays a role in whether time and resources are made available for data analysis, interpretation, planning and implementation (Schildkamp & Teddlie, 2008).

The beliefs about the intended use of the monitoring and feedback are of cardinal importance and if aimed at accountability they are often met with resistance. Monitoring aimed at improvement and problem-solving is more likely to be received positively and result in active use. During any monitoring and feedback process, people are also less resistant to change when the programme is less established, than later on in the process (Plomp, 2009). Some types of monitoring activities are also more likely than others to meet with a high level of resistance, for instance feedback that calls for change in a school entrenched in its activities and traditions. Conversely, resistance to such feedback in a young school that has a culture of change will be lower. The choice of monitoring activities may also have an impact on how monitoring is perceived and received. Teachers may perceive monitoring activities such as self-evaluations as less threatening than expert reviews or observations of classroom activities.

Certain characteristics, referred to here as ***relationship flux characteristics,*** are a product of the user-facilitator interaction and are therefore in constant flow, changing throughout the process. These characteristics are credibility, trust and sense of ownership. The perception of the credibility of the researcher is constantly being

evaluated and adjusted throughout the use of the monitoring and feedback, as is the user's sense of ownership of the process. The two-way trust relationship between the user and facilitator is also continuously being modified. Level of trust issues in monitoring and feedback systems are seen as a major determinant in the use of these systems (Taut & Alkin, 2003). All of these relationship flux characteristics are mediated by the monitoring characteristics, and dialogue throughout the process, including the sense of ownership that the users feel.

The ***feedback system*** referred to here encapsulates the idea of dialogue. Feedback often has the connotation of one-way communication that usually only takes place towards the end of a monitoring process to convey results. Nevo (2001) introduced the concept of dialogue instead of feedback in evaluation that highlights the interactive, two-way flow of information on a continuing basis. Originally, dialogue was introduced as an interaction between internal and external evaluation. In this thesis, feedback refers to a similar interaction between the user and researcher. Substantive issues that play a role in feedback include the quality of the communication products, the timeliness of reporting, communication, the accessibility and quality of the data, relevance and correspondence with the needs of the users (Alkin & Daillak, 1979; Alkin & Taut, 2003; Hattie, 2005; Johnson, 1998). The research paradigm has an indirect influence on the dialogue as it influences the roles associated with the users and facilitator. This in turn influences the type of communication, its quality and the level and extent of engagement between the user and facilitator.

Cousins and Leithwood (1986) conceptualised use of monitoring and feedback as a process, acknowledging use as both a decision and educational process that dynamically interact with each other. Process use and findings use are depicted as employing similar mechanisms in the conceptual framework.

Process use refers to use of monitoring and feedback as it takes place during the monitoring process, occurring when certain realisations form that may already lead to action and use prior to receiving the full findings. ***Findings use*** refers to use of monitoring and feedback of the results, recommendations and findings. Any type of use of feedback and monitoring commences with enlightenment. Owen (in Johnson, 1998, p. 103) referred to ‘enlightenment’ as the precursor of all other utilisation,

sometimes referred to in the literature as 'conceptual use'. Weiss (in Patton, 1997, p. 82) associates enlightenment with decision accretion, a process through which the user's existing knowledge is supplemented by new knowledge from the current monitoring and feedback. Further decisions are based on the user's entire body of knowledge on the phenomenon, and not just that obtained from the monitoring and feedback. Enlightenment and conceptual use also refer to learning about the process of monitoring, the methodologies and the ways of thinking and planning that may also change behaviour not specifically related to the monitoring and feedback. For instance, a school involved in the SAMP project my start using the same approach of data discussion and action planning used in the feedback sessions with other types of data from different sources.

Once enlightenment has taken place, the user combines his or her own body of knowledge with the competing information available in the external environment. This information is not all seen objectively; but is subjectively viewed through a lens coloured by the ***political demands*** of the context. This may include issues as diverse as financial limitations, the status of the user and external pressure for change from the DoE. For instance the DoE may decide to assess the functioning of a school purely on the Grade 1 performance on a mandated assessment with no consideration of the intake characteristics of the learners. Schools may therefore try to undermine the process by asking weaker learners to stay at home on the day of the testing in order to avoid any high-stakes consequences. This lens is so powerful that the knowledge gained may be distorted into misuses such a symbolic use, purposeful undermining of the process, legitimative use (the use of data to support and justify decisions already made that were not based on the data) (Owen in Alkin & Taut, 2003, p. 5) or purposeful non-use of the data.

If the ***political demands*** lens does not immediately distort the information into misuse, information is further ***filtered*** through reasonability and feasibility testing. This relates to the concepts of truth and utilities testing introduced by Weiss and Bucuvalas (in Patton, 1991, p. 291). ***Reasonability testing*** is common in accounting and mathematics and involves rounding and quick mental calculations in order to determine whether an answer is reasonable, or whether the calculation should be rechecked. In monitoring and feedback, this characteristic would relate to users

quickly examining results and information to establish whether these can reasonably be accommodated within the user's current understanding of the context and phenomenon. If the results seem too far removed from the user's current existing knowledge of the situation, he or she may well reject them. The term ***feasibility testing*** is preferred to that of utility testing, as the term utility seems to imply that use is purely determined by this assessment. A feasibility assessment examines whether the information gained from the monitoring could realistically be used to bring about change. If the results relate to something that cannot be changed, or which cannot reasonably be addressed with available resources, it may well be ignored and therefore not bring about any change. If the body of knowledge available to the user combined with knowledge gained from the monitoring and feedback fails either one of these tests, misuse may occur through an undermining of the process or purposeful non-use of the data. It is only if the reasonability and feasibility testing are passed that instrumental use can take place, both during the process, when interim data are used for action, or as a result of the findings reported towards the end of the monitoring and feedback. The term instrumental use was first introduced in 1977 by Rich (in Johnson, 1998, p. 93) and refers to direct action based on data.

This cycle of use does not conclude with the use itself, but there is an influence from the use on the formation of the monitoring culture in which the user functions. Throughout the user's perception of monitoring feedback and prior experiences of monitoring and feedback are changed according to their current experiences of the feedback and monitoring as they take place. This continuous process means that enlightenment or conceptual use that takes place may influence the user's amenability towards use of the findings of the same monitoring and feedback process at a later stage in the interaction.

The concept of ***distal effects*** relates to Kirkhart's (in Alkin & Taut, 2003, pp. 8-9) idea of influence that extends beyond the monitoring and feedback. This is also sometimes referred to as distal outcomes (Rossi, et al., 2004) and they take place well beyond the bounds of monitoring and feedback. The level of awareness and intention about these effects may vary drastically from one feedback and monitoring process to another. Distal effects are unpredicted and may present themselves in aspects of the school very different from those the feedback was originally intended

to influence. Distal effects are beyond the scope of this study, but do form part of the full conceptualisation of use of monitoring feedback.

In this research that utilises the SAMP project, external school-based monitoring is used. Staff of the CEA administer the SAMP assessments, with the data analysed at the CEA and the CEA feeding reports back to the schools. This process is moving towards becoming more internally driven. The long-term goal is for the SAMP system to be administered and managed by schools themselves. Participating schools will be provided with ICT tools to analyse the data by themselves. Currently, as this research is still using an external monitoring system, the work done on use of feedback in this thesis may need to be extended to see how it applies to an internal monitoring process. Although this conceptual framework tries to tease out and clarify the complexity of feedback use, there are certain levels of complexity inherent to systems theory that have to be acknowledged.

3.5.3 The Complexity of Change

It is essential to note that monitoring and feedback are not conceptualised as some black box, or that only the processes and findings are of interest in this study. The components illustrated in the conceptual framework can vary vastly and lead to diverse change in the school. A highly rigorous monitoring process can take place leading to valid and relevant findings that are fed back to the schools. These findings and processes may then either be misused by the school or ignored. It is also possible that poor monitoring and feedback will generate findings that are used by the school and leads to positive change. In the same way, a principal or other user may feel that findings of rigorous and valid monitoring may lead to despondency and so not contribute to positive change in the school. The principal may then withhold the information from his educators, while personally driving for directed change using different avenues. In such a case, positive change may occur through the purposive non-use of findings. Complexity of change is closely related to the issue of use and misuse.

Some authors such as Christie and Alkin (1999) see use and misuse as constituting completely different continua and seek to track the courses of actions inherent in various types of misuse. This may be misleading, as use and misuse of monitoring sometimes only shows a tenuous link to the type of change that takes place. Weiss

(in Shulha & Cousins, 1997, p. 202) warns against becoming comfortable with a definitive framework for understanding misuse by saying that "…when you start taking into account the larger social good, it becomes very hard to make judgements about the immediate use and misuse of evaluation. People do rather wondrous things because there is a larger good involved." This becomes a slippery slope of the 'ends justify the means', as even misuse may result in positive change in the system, whereas conventional appropriate instrumental use may well result in adverse change.

Nevo (2007) records a case of how over-interpretation of the relative ranking of Israel in comparative international studies has led to positive change. During the period 1999-2002 Israel's position on the international comparative studies such as TIMSS, PIRLS and the Programme of International Student Assessment (PISA) dropped. The drop was not particularly significant, since the sample for the different studies changed with the introduction of the Arab population to the Israeli sample. Newly participating countries were also introduced into the studies, which also affected country rankings. The results, however, sparked new interest in education and the quality of education in the Israeli government, which led to initiatives to improve education across the country. While various change outcomes can result from taking different routes, rigorous, relevant and valid monitoring should be striven for at all times, to drive for directed positive change in teaching and learning.

All systems as in this conceptual framework also inherently strive towards a state of homoeostasis (stability and balance) and as such tend to resist change. This resistance to change may to some extent explain the lack of use of monitoring and feedback. For Fullan (2006), the key to overcoming this resistance to change is motivating the people in the system, thus working on numerous parts of the system at once. The principle of ***equifinality*** is also important, that is when changes are introduced into a system, in this case in the form monitoring and feedback, they are not necessarily predictable due to the complexity of the various interacting factors, conditions and the complex feedback loops in the system (Voster, 2003). One such feedback loop is illustrated in this conceptual framework where the monitoring and feedback culture of the users is constantly being modified by the experience and use of the feedback system.

3.6 Conclusion

A multitude of SPFSs, monitoring and evaluation systems exist in the world, most of which have similar characteristics, employed to help provide the design criteria for this feedback system. The literature on monitoring, evaluation, feedback, school improvement and school effectiveness are employed to create a conceptual framework for this thesis. Change through use of feedback from monitoring is a complex process. The process is influenced by a multitude of characteristics and dynamics. Data are weighed against other available data and the information is assessed in terms of the reasonability and feasibility for change before use takes place. Even if rigorous data are generated and fed back to schools, use is still influenced by environmental factors and user perceptions. In the next chapter, the methodology employed in this thesis is examined.

CHAPTER FOUR

Overview of the Research Design

> …[P]ragmatist researchers have the opportunity to combine the macro and micro levels of a research issue… [T]hey are more likely to be cognizant of all available research techniques and to select methods with respect to their values for addressing the underlying research questions, rather than with regard to some preconceived biases about which paradigm is a hegemony in social science research. (Onwuegbuzie & Leech, 2005, p. 291)

This chapter discusses the research design and methods employed to answer the research questions. The researcher considers how the research question informed the paradigm and the research design and explores why the chosen research design was the most appropriate to address the research question. How the design research process was applied to address the specific research questions is also discussed. Each design research cycle presents a small but complete research process that leads to findings that inform the next cycle of development.

This chapter also serves to introduce the design research process as so few design research studies have been published in peer-reviewed forums (Barab & Squire, 2004; Collins, Joseph, & Bielaczyc, 2004; Kelly, 2003; Nieveen, 1997; Plomp, Educational Design Research: an Introduction, 2009; Reeves, Herrington, & Oliver, 2005). The chapter therefore provides additional details on design research to make this study more accessible to readers who are unfamiliar with this specific research design. The methodology, research procedures and data for each cycle are discussed separately and in more detail in Chapters 5-7. The research design is also discussed in terms of the evaluative criteria that were applied. Finally, the methodological norms and ethical research behaviour maintained through this study are considered.

4.1 Research Paradigm

The primary research question in this thesis is: “What are the characteristics of an optimum feedback system to facilitate appropriate use of learner performance monitoring in primary schools in South Africa?” The research aimed to optimise a

feedback system for a learner performance monitoring system. The design research methods produced both an optimised feedback system and design principles that contribute to the body of knowledge on feedback systems.

This design research process was located in the pragmatist paradigm, which foregrounds the need to answer the research question comprehensively and appropriately above paradigmatic purism. Therefore, the most appropriate tools to answer the research question are employed, irrespective of whether the methods were qualitative, quantitative or a mixture of both (Onwuegbuzie & Johnson, 2004, 2006; Onwuegbuzie & Leech, 2005; Tashakkori & Teddlie, 1998).

A paradigm can be described as "[t]he underlying philosophy and assumptions that form the foundation to one's approach and methodology" (De Villiers, 2005, p. 120). The issue of choice of paradigm is much contended and often polarises quantitative and qualitative researchers. This polarising resulted in an extended conflict amongst purists known as the paradigm wars (Tashakkori & Teddlie, 1998). In the midst of the paradigm conflict some researchers lost sight of the function of research methodologies as tools to facilitate understanding of phenomena. This lack of insight gave rise to religious support of specific philosophical underpinnings (Onwuegbuzie, 2002). The third wave or pragmatist paradigm emerged as the culmination of this conflict.

Researchers working in the pragmatist paradigm reject the research paradigm as the starting point for research and focus on what research design is able to address the research questions (Greene, 2007; Mertens, 2009; Tashakkori & Teddlie, 1998). Pragmatists willingly combine multiple, diverse measures to address the research question appropriately, an approach that often combines qualitative and quantitative methods. This paradigm aims to employ the complementary strengths of various methods to supplement the weaknesses of other methods. Research questions can usually be answered in this way more appropriately than with a method that applies solely qualitative or quantitative methods (Onwuegbuzie & Johnson, 2004).

Paradigms are usually discussed in terms of the basic underlying beliefs or philosophy underpinning the paradigm. These usually include:

- ontology: assumptions about the nature of what exists and what is viewed as reality
- epistemology: assumptions about the nature of knowledge and the relationship between researcher/evaluator and the stakeholders needed to achieve accurate knowledge
- axiology: assumptions about ethics and values
- methodology: assumptions about appropriate methods of systematic inquiry

(Mertens, 2009, p. 49)

However, with the pragmatist paradigm, each of the multiple methods employed are grounded in the philosophy and tenets usually associated with the particular method. The research during any given cycle upheld the tenets and axioms of either the social constructivist or post-positivist paradigm, depending on the methodologies employed[21]. Although this characteristic complicates discussion of the tenets of the pragmatist paradigm, the next sections examines these tenets as applied in this study.

4.1.1 Ontology

Pragmatists' embrace of ontological pluralism obliges them to accept both objective and subjective realities as:

> …a very broad and inclusive ontological realism where virtually everything a qualitative or quantitative researcher deems to be real can be considered, in some sense, to be real, including subjective realism, intersubjective realism, and objective realism. (Onwuegbuzie & Johnson, 2006, p. 54)

The qualitative methods in this study were located in the social constructivist philosophy. Ontologically, social constructivists see realities as constructed through human interaction. Multiple realities exist and are constructed, interpreted and subjective. (Gergen, Lightfoot, & Sydow, 2004; Maddux, Snyder, & Lopez, 2004)

When quantitative methods were employed, the research was grounded in the post-positivist paradigm, which retains the positivist ontology of one reality that exists independently of knowledge and perception (Greene, 2007). However, post-

[21] The different foundations of these paradigms had implications for the methodological norms of this study, which are discussed in more detail in Section 4.3.

positivists acknowledge human fallibility as observers, in this way moving beyond the naïve realism of traditional positivism.

4.1.2 Epistemology

A Pragmatist believes in the existence of causal relationships, but acknowledges that these cannot always be determined definitively (Tashakkori & Teddlie, 1998). The social constuctivist epistemology that formed the basis of the qualitative methods purports that events are interpreted to be understood and social context mediates this process. Meaning and knowledge are created or constructed, allowing people to make sense of their worlds (Gergen, et al., 2004; Maddux, et al., 2004).

The social constructivist paradigm sensitised the researcher to the influence of the evaluation cultures in schools, individual personality factors and relationships on the use of the feedback system. The paradigm helped in the exploration of the constructed and interpreted factors that form the users' multiple realities concerning the feedback system. These realities may not always be overt, but play an important role in sustainable and appropriate use of feedback and monitoring.

> When I work with the reams of data from the schools, the questionnaire data, learner data, Rasch analysis and classical test theory, I often feel disconnected from the research. It can be so clean and clinical. But I often feel that it is only scratching the surface. The numbers hide the complexity of emotion and human interaction, which I cannot help but feel is a major driver in this feedback system. I don't have relationships with schools, but individual persons, each person has their own fears and needs and see me as facilitator in a different way. It feels good to sometimes pick up the phone and make sense of the numbers by speaking to a principal or reviewing some of the raw interview transcripts to get back to grips with the human element of the feedback system. (Research diary 13/11/2008)

Epistemologically, post-positivist approach that grounded the quantitative methods still aims to establish objective truth that leads to "generalisable causal explanations of observed human phenomenon (sic)" (Greene, 2007, p. 38). However, the fallibility of the researcher is acknowledged. Post-positivists believe that standardised quantitative instruments decrease such researcher subjectivity and bias (Mertens, 2009).

4.1.3 Axiology

The tenets of social constructivism allow the researcher to use inductive analysis of interviews, observations and reflection data. The role of construction is fore grounded not only in terms of the construction of instruments and generation of data, but also as far as the analysis and interpretation is concerned. Therefore, the impact of the researcher's dual role as monitoring facilitator and researcher is acknowledged and examined.

> …[A]ny analysis is contextually situated in time, place, culture, and situation. Because constructivists see facts and values as linked, they acknowledge that what they see – and do not see – rest on values. Thus, constructivists attempt to become aware of their presuppositions and to grapple with how they affect the research. (Charmaz, 2006, p. 131)

Ethically a social constructivist is bound to reflective practices such as keeping a research diary, peer review and member checking (see Section 4.3 for a full discussion). Schools were not seen as mere participants, but rather as research collaborators and as experts on their own context who could take an active role in the research process.

The post-positivists believe that standardised quantitative instruments help decrease researcher subjectivity and bias, helping to balance the social constructivism. With this paradigm, aspects of the use of feedback system could be quantified and weighed as seen with the Delphi technique and questionnaire data (see Chapter 6).

> I have been analysing the interview data for a while. Each interview is rich with contextual factors and emotions. Sometimes it seems that the contexts, cultures, leadership approach etc. of the schools differ so much that there are little similarities. The quantitative data give me a sense that there are some golden threads that run through. Common needs, preferences and concerns which can be quantified, if only in a rudimentary fashion (Research diary 02/10/2009)

The ethical and value implications for this study are discussed in greater depth in Section 4.3.

4.1.4 Methodology

In order to address the research question most appropriately, this study capitalised on the complementary strengths of both qualitative and quantitative methods through triangulation of method. According to Onwuegbuzie and Johnson (2004, p. 17) the mixing of method allows for "…the use of induction (or discovery of patterns), deduction (testing of theories and hypotheses), and abduction (uncovering and relying on the best of a set of explanations for understanding one's results)".

Inductive and deductive logic served to optimise the feedback system. Initially the global design elements for the feedback system were established, followed by the detailed development and evaluation of the identified components. The inductive logic inherent in the qualitative methodologies was well suited to exploring the global design elements and informing the overall design of the feedback system. The practicality and current declared use of the feedback system by and in schools were also explored inductively. The qualitative methods allowed for exploration of the rich experiences of users of the feedback system. Quantitative methods were employed to explore the extent and type of use taking place before a rich qualitative exploration was conducted.

The deductive logic of the quantitative methodologies was used to design and evaluate the quality of the specific components of the feedback system, such as the reports and feedback sessions. Each particular aspect of the feedback system under development during each design cycle could be targeted and evaluated in this fashion. The quantitative methods also made it possible to gather input from all the primary users of the feedback system. Input from a greater number of primary users could be gathered through the quantitative methods. When users and experts evaluated the design quantitatively, weighted design priorities could be produced to guide the design process.

The research design for this thesis was design research. The following section first discusses the design research approach in general, then explores how design research was employed to address the research questions in this study.

4.2 Research Design

The study applied design research (also called development research) as the most appropriate design to address the research questions as fully as possible.

> Development research is problem-oriented, searching for new and innovative solutions, while also seeking findings that are transferable, practical, and socially responsible. [Development research] acknowledges the complex and dynamic relationship between theory and application, and aims to provide a relevant foundation to guide practice by generating design principles and methods that are both theoretically underpinned and empirically tested. (De Villiers, 2005, p. 114)

Sub-question 1, relating to the adaptation, translation and contextualisation of a performance monitoring system, is addressed in Chapter 2 by means of teacher evaluations, classical test theory, Rasch modelling and an expert panel review. The characteristics documented in literature for an optimal feedback system, sub-question 2, are explored in Chapter 3 through a literature review. Sub-question 3 deals with the pre-existing conditions that need to be established for a feedback system to facilitate use. This is addressed in chapters 5-6 by means of data generated through expert evaluations, the Delphi technique and questionnaires. Sub-question 4 examines how schools use data. This is addressed in Chapters 6-8 employing data from observations, reflective journals and interviews. The efficacy of the feedback system is the focus of sub-question 5 and is examined through the final evaluation questionnaires and expert evaluator reports in the semi-summative assessment documented in Chapter 8. The design guidelines of sub-question 6 are the product of the entire design research process and together with the main research question are addressed in Chapter 9. These links between the methods and research questions are graphically illustrated in Figure 4.5 and summarised in Table 4.2.

The application-orientated philosophy of design research is congruent with the aims of this study, as the focus was on designing and optimising a feedback system and then identifying its characteristics in particular those characteristics that facilitates use. Design research aims to:

- provide more solid, timely and appropriate data for people working in educational improvement than traditional research offers

- allow for more iterative, cyclical processes to integrating research into educational reform
- allow educational researchers to have a more direct impact on educational reform

(Van den Akker, 1999, p. 2)

Design research starts with the design or improvement of an artefact or intervention to address a real world problem. In this study, the feedback system for a learner performance monitoring system was optimised. As this feedback system is based on and intricately linked to the SAMP monitoring system, the various feedback system prototypes[22] extended and also influenced the development of the SAMP monitoring system itself.

4.2.1 Design Research

The purpose of design research is to blend design and research. Design research is conducted to understand the issues of application, not only theory (De Villiers, 2005; McKenney, 2001; Reeves, et al., 2005). The aim of design research is therefore to address problems in practice for which no guidelines for design solutions are available.

In essence, design research is a cyclical iterative analysis of design, development and implementation; formative evaluations assess each completed cycle and inform subsequent cycles of design. Van den Akker (1999, p. 2) sums up this process as the "successive approximation of the ideal".

Throughout the design process the focus of the cycles shift, a progression that moves through three distinct phases:

- ***Preliminary Phase:*** This addresses the needs and context analysis, which includes a literature review to conceptualise the intervention or the product and defines the design specifications.
- ***Prototyping Phase:*** This consists of the iterative, cyclical research cycles during which various prototypes of the product or intervention are fashioned

[22] A prototype in this case is a version of the feedback system that includes the design, implementation, evaluation and adaptation processes. The version includes the reports, feedback sessions, electronic resources and support for transforming the monitoring into school improvement. Each version is formatively evaluated and the evaluation provides the basis for development of the next version or prototype.

and then formatively evaluated to inform the development of the next prototype.

- ***Assessment Phase:*** This is the summative evaluation[23] that examines the functioning of the elements of the product, as well as the global functioning of the integrated system. It is the final phase of the design research process, but may be followed by further design research aimed at 'up-scaling'.

  (Plomp, 2009)

Design research not only progresses through various phases, but also has particular underlying features and characteristics. Wademan (in Plomp, 2009, p. 16) graphically illustrates these features in Figure 4.1. The model depicts the iterative, cyclical nature of design research and the integrated development of product and theory. It highlights the active involvement of practitioners, experts and researchers in the various phases of design and development.

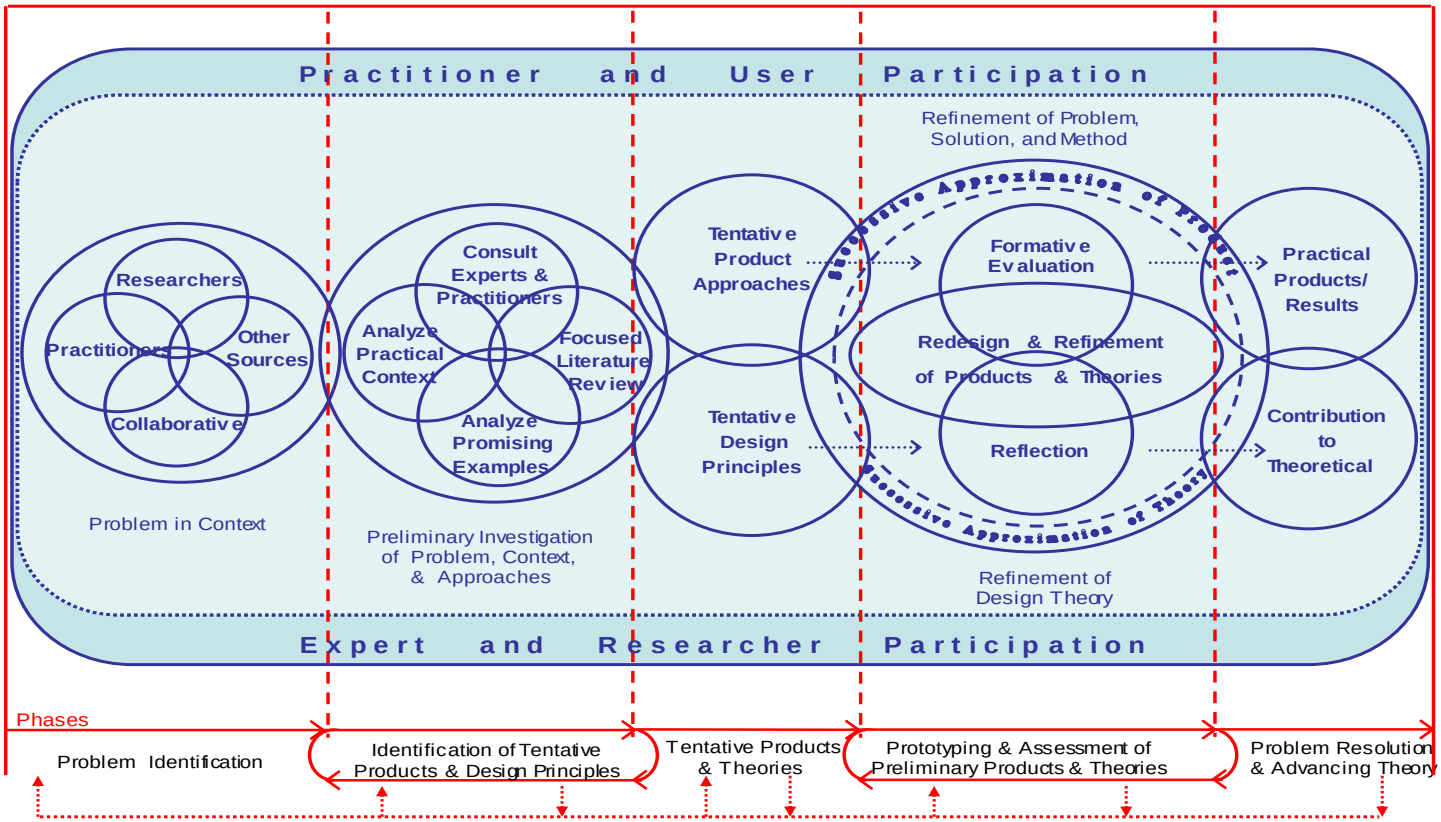


Figure 4.1: Wademan's Generic Design Research Model

[23] In this thesis, the evaluation is semi-summative as only the perceived efficacy of the feedback intervention was examined. The feedback system will be further developed as part of the CEA project during 2010, with a focus on actual efficacy.

Not only are there multiple cycles throughout the application of design research, but each cycle also encompasses a series of activities:

1. ***Design and development*** – the data from the previous cycle's evaluation are interpreted and informs planning and design of the next prototype.
2. ***Implementation*** – involves testing of the new prototype. Implementation may involve a trial of the adjusted prototype in the field, or may be the presentation of planned changes and a mock prototype that is not necessarily tested in the field, but can be evaluated in principle by users and experts.
3. ***Evaluation*** – encompasses judgement by experts and users about the new prototype. Feedback is not only critical of errors, but also aimed at providing a basis for further development and improvement.

These activities take place within each of the design focus phases noted above (Plomp, 2009).

During design research, users and experts formatively evaluate each prototype of the feedback system. The evaluations can be informal to formal[24]. Tessmer (in Plomp, 2009, p. 28) created a model of the layers of formative evaluation (see Figure 4.2). In this model, Tessmer shows the continuum of resistance to revision that may be expected during various levels of formative evaluation in a development process. From this illustration, it becomes clear that during any formative evaluation process, resistance to change is lower early in the development process, with resistance increasing as the prototype becomes more refined and established. Therefore, more informal formative evaluative activities such as self-evaluations and expert reviews are often used towards the start of revision while more formal methods, such as field tests, are usually applied later on in the process when resistance to revision is higher.

[24] Although the evaluations may be informal, there need to be guidelines and foci for each evaluation and it should be documented appropriately as part of a rigorous research process.

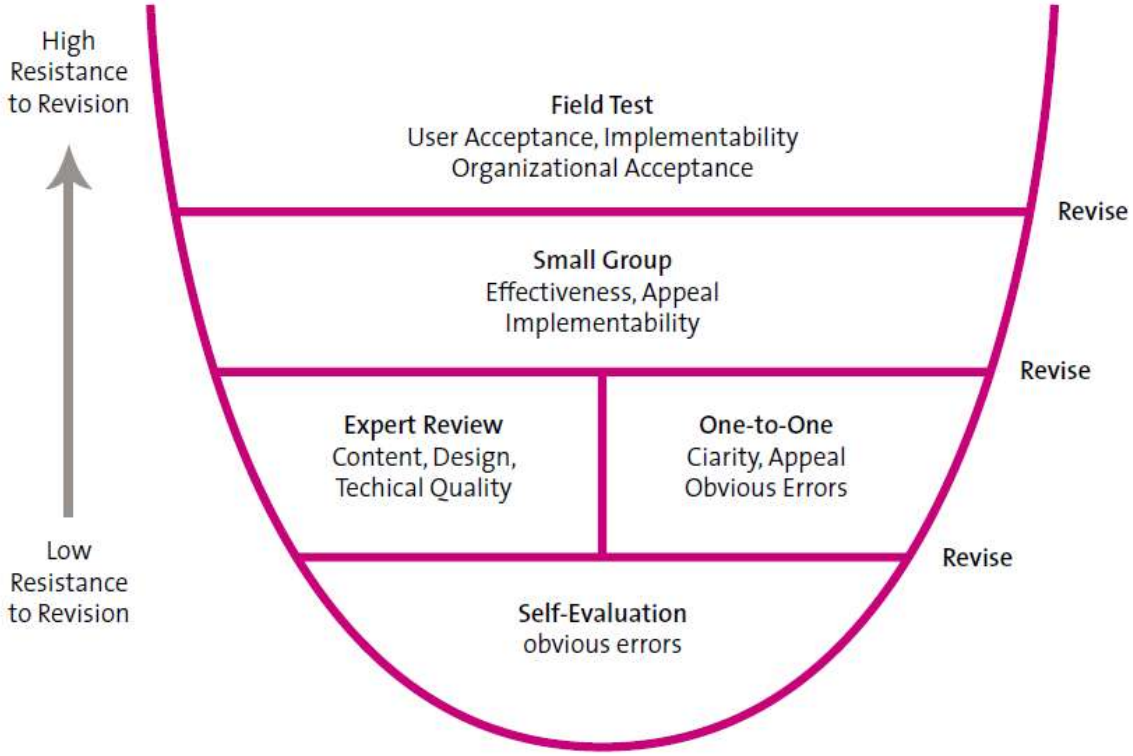


Figure 4.2: Layers of formative evaluation
(Tessmer in Plomp, 2009, p. 28)

In the design research process, the data from the evaluations are immediately used to further develop and adapt the next prototype. This new prototype is then tried out and evaluated in turn. In this way, there is movement between evaluation and design in order to produce the best possible approximation of the ideal. The question arises as to when the design cycles cease. In short, the answer is: when the design fulfils the necessary evaluative criteria or when cost-benefit analysis dictates it.

4.2.2 Evaluative Criteria in Design Research

The formative evaluations of design research are systematic activities performed to achieve well-directed improvements to prototypes. The purpose of this investigation was to design a high quality feedback system and identify the associated design principles. The evaluations helped achieve insight into what was required to design an ideal feedback system and to evaluate how close to realising this ideal any of the prototypes were. Evaluations are by definition judgements about worth or merit (Joint Committee on Standards for Educational Evaluation, 1994) and it is necessary to discuss the criteria used to evaluate this design research process and product.

Nieveen (2009) explains that a completed design intervention should fulfil four quality criteria (see Table 4.1), these four quality criteria are hierarchical. The intervention cannot be ***effective*** or have catalytic validity unless it is ***practical*** in the setting for which it was designed. ***Consistency***, also referred to as 'construct validity' is established with reference to ***content validity*** (relevance).

Table 4.1: Quality criteria for the feedback system

CRITERION	DESCRIPTION
Relevance (content validity)	There is a need for the system and its design is based on state-of-the-art (scientific) knowledge. In addition, the feedback system must be clearly connected, in form and purpose, to the learning performance monitoring system for which it provides feedback.
Consistency (construct validity)	The system is 'logically' well-designed. The various parts are well defined and the connections between the parts explicitly postulated. Although some elements may emerge more clearly throughout the process, there are no internal contradictions.
Practicality	**Expected:** The system is expected to be usable in the settings for which it has been designed and developed. **Actual:** The system is usable in the settings for which it has been designed and developed.
Effectiveness (catalytic validity)	**Expected:** Using the system is expected to result in desired outcomes. **Actual:** Using the system results in desired outcomes.

(Adapted from Nieveen, 2009, p. 94)

These four criteria were not all examined in every design cycle. As the design progresses from the Preliminary Phase through the Prototyping Phase into the final Assessment Phase, the evaluative emphasis also shifts through the hierarchy of criteria. Usually each cycle focuses on one or two criteria to evaluate a current feedback system prototype. In this way, relevance and consistency were acutely attended to early in the process. Practicality and effectiveness are of major importance during the later design cycles. The shift in design phase and evaluative focus of the cycles is illustrated in Figure 4.3.

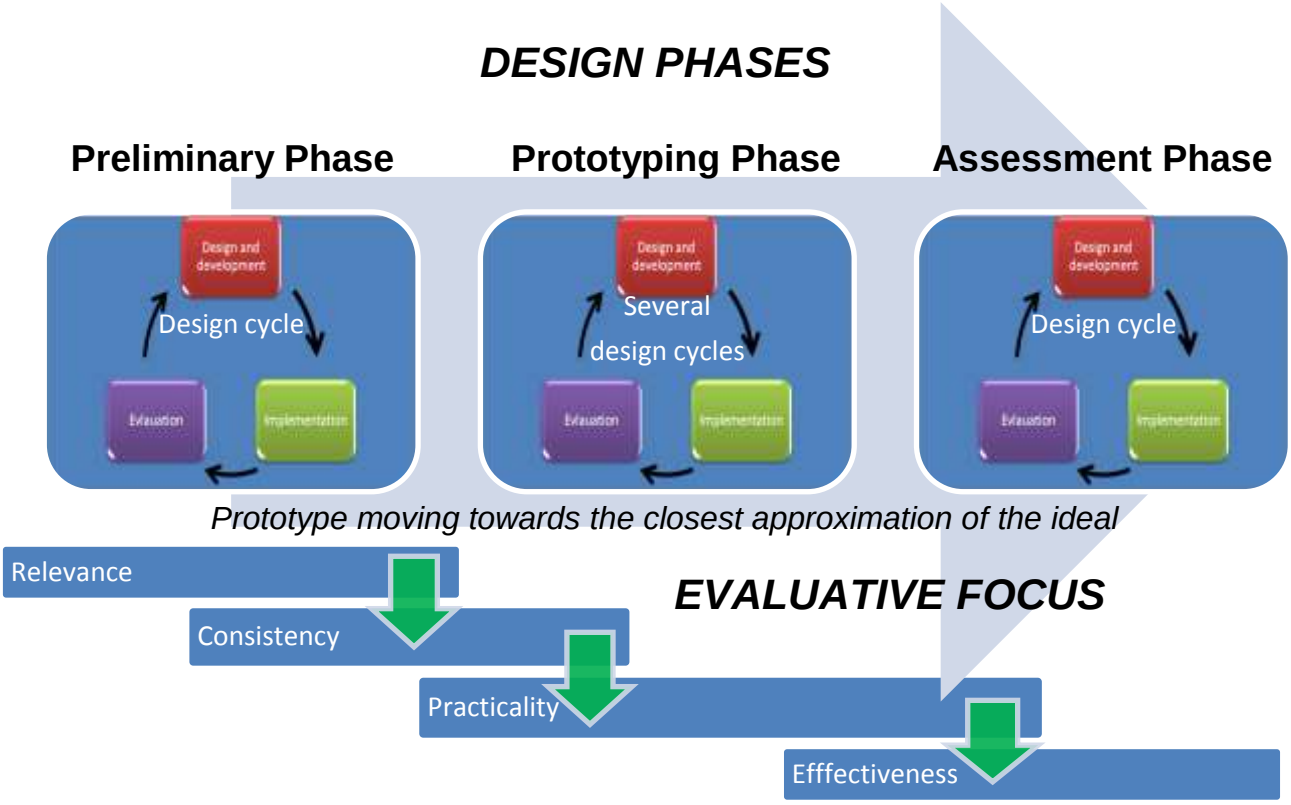


Figure 4.3: Interaction between design phase, design activities and evaluative focus in Design research

Whilst applying the evaluative criteria during the various cycles, the different participants in the design research evaluations may also take on a variety of roles. Nieveen (2009) differentiates three types of roles for participants in design research evaluations:

- **Learner:** This respondent is not an expert in the specific subject matter. Some learning is required on the part of the evaluator before he or she can bring his or her expertise to bear. The academic experts in this research in certain respects took on the role of learner. Some had the expertise in the methodology, but were not familiar with the South African context, while others had expert knowledge on utilisation of monitoring, but needed to develop their knowledge of design research. School user participants had extensive knowledge of their contexts, but needed to learn something of the design research process and goals to become full collaborators.
- **Critics:** Participants comment on the intervention from their fields of expertise. The role of critic was the main role taken by the school user participants in data collection and was accomplished through questionnaires

where critics were asked to comment on reports, feedback sessions and support.

- **Revisors[25]:** Participants as revisors not only comment on the intervention, but were also asked to provide recommendations for improvement. This type of input was required of academic expert evaluators during all phases of evaluation. Specific questions aimed at eliciting suggestions for improvement were also presented to the school users, throughout the research.

The design research process is only complete when the evaluative criteria are fulfilled and the solution or product, including its accompanying design theory, is generated. Table 4.2 (at the end of this chapter), illustrates how the evaluative criteria for this thesis were met together with how the evaluative focus shifted through the design phases. The emphasis on producing an accompanying design theory or principles separates pure design and development from design research. The developed design theory aims to support other practitioners and researchers working on similar issues. In this way design research contributes to the existing body of knowledge on optimising a learner performance monitoring feedback system.

The following section examines how this design research approach was applied to the design and optimisation of the feedback system. The overall design research process is examined, detailed research procedures for each cycle are discussed separately in Chapters 5-8.

4.2.3 Application of Design Research for this Inquiry

The SAMP system provides the learner performance data for which this feedback system was optimised. The SAMP system has been in development since 2003 and incorporated a basic feedback system prior to this study. The basic feedback system provided some data to schools from the monitoring system, but schools found it difficult to interpret the feedback and transform it into improvement action. This need expressed by schools resulted in this study. It is hoped that the principles and design

[25]The term 'revisor' as opposed to the British English spelling of 'reviser' is used throughout this text as it is the term applied to this role by Nieveen (2009).

theory for this study have potential application beyond the scope of the SAMP monitoring system to other feedback systems in educational settings.

Each design cycle constitutes a complete micro-cycle of research on its own and may incorporate a variety of methods, both quantitative and qualitative. Each cycle requires users who implement the latest design prototype in the particular context, but also experts and users who act as evaluators. How the sampling takes place may change from one cycle to the next. In the design process, smaller groups may be asked to implement the prototype, while a larger group may be involved as the prototypes approach the ideal. The evaluative focus for a cycle may also influence whether or how many experts or users are required to provide input on a particular cycle.

The research and design process for this thesis is first shown overall to indicate how the cycles progressed through time and were built on one another. The progress of the cycles is depicted in Figure 4.4, gives an overview of both the progression and the different populations used for implementation and evaluation during each complete design cycle.

Each cycle in the Prototyping Phase consisted of the design and development of a prototype of the feedback system, followed by an evaluation designed to address a particular research sub-questions and evaluative criteria (McKenney, 2001; Nieveen, 1997; Richey, et al, 1996; Thijs, 1999). Figure 4.4 (Based on the representational style of Nieveen, 1997) illustrates the iterative and cyclical pathway of design research process along with the evaluative criteria addressed in each particular cycle for this thesis. The samples of participants in the *Implementation* and *Evaluation* of the different feedback system prototypes are indicated on the Y-axis (vertical). The number of participants shown above the X-axis (horizontal) refers to the teachers, HoDs and principals who received information from the SAMP monitoring system through the feedback system.

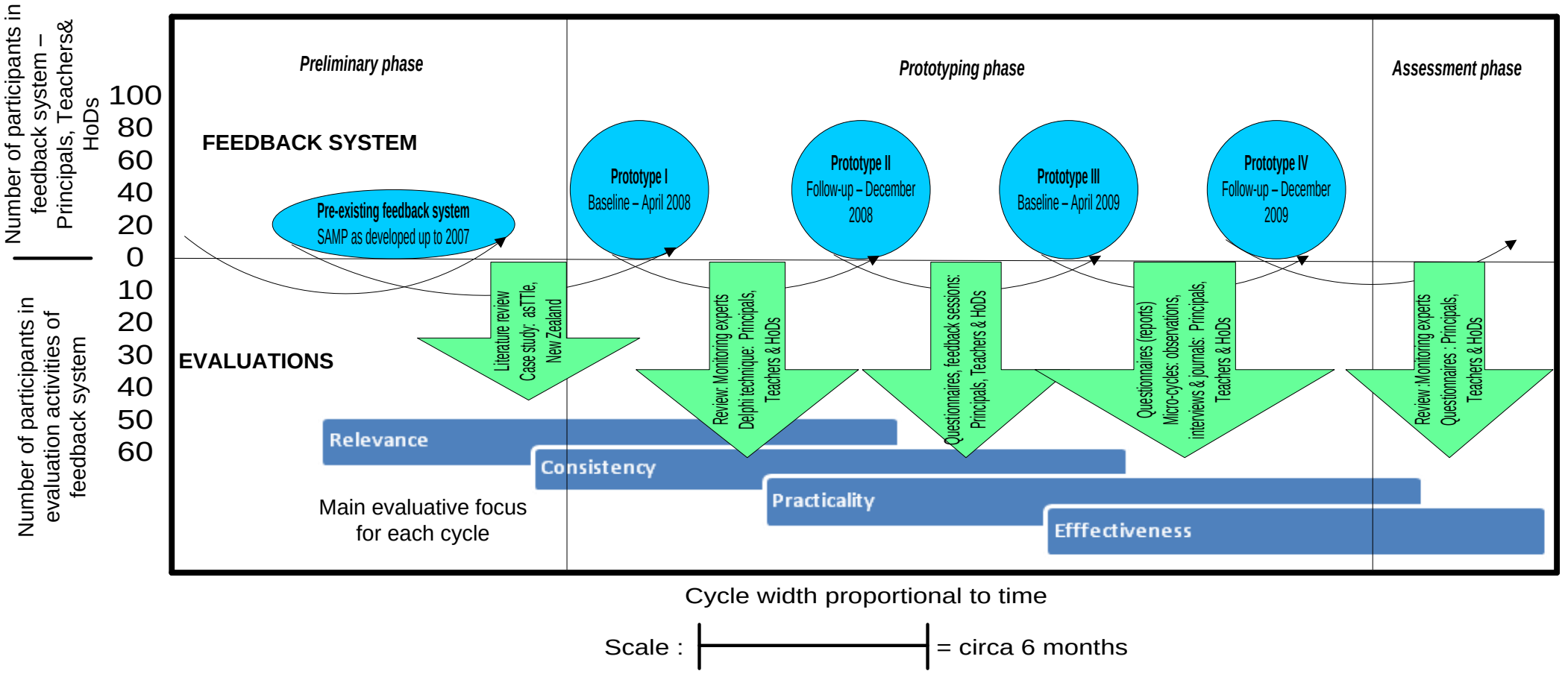


Figure 4.4: Design research cycles for this thesis

The feedback prototypes report on the baseline and follow-up assessment of learners in the SAMP project. The four prototypes (I-IV) served as the feedback system for the learners' performance data from the SAMP project. Each feedback system prototype was evaluated focussing on the evaluative criteria indicated in blue, below the X-axis (horizontal axis). The evaluation in Cycle 4 was also slightly different as the evaluation was iterative and consisted of a number of evaluation activities that built on each other. This thesis moved through five complete cycles, but development will continue as part of the work of the CEA.

The X-axis in Figure 4.4 indicates the time element in the cycles. The pre-existing SAMP monitoring system was the culmination of work at the CEA initiated in 2003. As the feedback prototypes coincide with the SAMP learner monitoring cycles, each cycle spans about 6 months. Each six-month period includes the assessment of the learners with SAMP, cleaning and analysis of data, production of reports, feedback sessions, interactions with schools and development of intervention materials to support schools.

Furthermore, the green arrows in Figure 4.4 represent the evaluations that took place as part of each full design cycle. The evaluations informed the tentative design principles and theories that influenced the development of the next feedback prototype. The vertical axis indicates the sample size of users and experts for each evaluation. Participants in the feedback systems are indicated above the horizontal axis[26]. User and monitoring experts who participated in the evaluation of each prototype are indicated below the horizontal axis.

Different categories of participants played a role in the evaluation, ranging from teachers, HoDs and principals to teacher trainers, experts in monitoring and academics in education. The final evaluation was only semi-summative as the design and development process of feedback system was set to continue after the conclusion of this thesis.

[26] Generally some 90 principals, educators and HoDs participated in the feedback system prototypes per cycle. Around 1,600 learners participated in each monitoring event that generated the data for the feedback system.

The formative evaluations for this research were not separate events that merely examined each successive prototype, but rather were linked and followed a design logic. The initial focus of the evaluations was on the global design of the product, after which attention shifted to how individual components of the feedback system functioned. Finally, the overall functioning of the feedback system was examined. The emphasis throughout the evaluations shifted from relevance and consistency to effectiveness.

The distinct developmental stages of the design research process for this feedback system are discussed below:

1. ***Design specifications and global design:*** These issues were addressed in the initial part of the process that explored the components necessary for the feedback system. A preliminary framework for the interaction of these aspects was also developed. The identifying and conceptualising was accomplished through the literature review, exemplary case study and previous work done on the SAMP feedback system. These activities formed the basis for the development of Prototype I.
2. ***Establishing conditions for use:*** This development stage was aimed at fleshing out the feedback components of the system, e.g. examining the design of reports and feedback sessions in detail. This aim was the main focus of formative evaluations of Prototypes I and II.
3. ***Transforming conditions for use into action in schools:*** This phase followed schools processes in employing the feedback system on a management, planning and classroom practice level. The focus was on the application of the report and feedback sessions data in the schools, along with the use of support materials, instrument manuals and linked intervention materials. The aim was to formulate a perspective on how these aspects were being used in the context for which they were designed and to inform further improvement of the intervention. Only expressed or presented use and limited instances of actual use could be studied for this thesis. This exploration was the focus of the micro-cycles of evaluation of Prototype III.

4. ***Completed intervention:*** The feedback system and its functioning as a whole was examined through the semi-summative evaluation of Prototype IV. In this evaluation, all components were evaluated along with the interaction of the components and their functioning as a whole.

Figure 4.4 also illustrated how the various prototypes were interspersed with formative evaluations to guide the next feedback prototype. There was an iterative process with cycles representing successive approximations of ideals. The details of the various design cycles and research procedures are discussed in the Chapters 5-8. The overall research procedures are discussed in Section 4.2.5.

The schools in the SAMP project represent the population for the feedback system optimisation. All the participating schools were sampled in a number of combinations during the design research process for the different cycles. For instance, all schools were invited to participate in the Delphi technique and questionnaires, but only a sub-sample were approached for the observations, interviews and reflective journals during Cycle three of the Prototyping Phase. The specifics of the sampling for each cycle are discussed separately in Chapters 5-8. The population for the feedback system is clarified below.

4.2.4 Population for the Design Research

The sample of schools for the SAMP project represents the population for this research, with Sepedi, Afrikaans or English Language of Learning and Teaching[27] (LOLT) public schools in the Tshwane region selected. These three languages were purposefully selected as they represent the three dominant LOLT in the Tshwane region (H. Julies, DoE, personal communication, July 16, 2010). This sampling resulted in the selection of 22 schools from three language groups: seven Sepedi LOLT schools, eight English LOLT schools, six Afrikaans schools and one dual LOLT English/Afrikaans school. This study was not aimed at statistical generalisability. The design guidelines and feedback system itself are intended to be adapted and applied in an appropriate manner to various contexts

[27] LOLT refers to the language in which the learners receive their education. In South Africa, learners often do not receive schooling in their home language.

through a process of analytical generalisability (as discussed in Section 4.2.6) The sample for SAMP therefore represents the accessible population from which the sample was drawn (Best & Kahn, 2006) for the SAMP study and therefore the feedback intervention.

Multi-phase sampling took place (Cohen, Manion, & Morrison, 2000) from the accessible population from SAMP. Firstly, schools were stratified according to LOLT. Secondly, eight schools were randomly selected from each LOLT from the DoE databases. Thirdly, the sample was inspected at this stage to ensure geographic representation of the Tshwane area to ensure maximum variation. The sample was found to be satisfactory in that regard as all the major geographic regions in Tshwane were adequately represented. This sample was maintained from 2007 to 2009 with only minor changes as in the case where some school characteristics changed e.g. one Afrikaans school changed to a dual-medium (Afrikaans/English) school during the period. In another case, an Afrikaans school withdrew from the study and was replaced by a school from the same area with a similar learner population in terms of home language, racial background and socio-economic status.

The population for the implementation and evaluation of the feedback system thus consisted of these 22 schools, as funding would not allow for expansion of the sample. All schools contributed to the evaluative activities of the five cycles of design research, but different schools from this population were sampled throughout the design research to evaluate the prototypes. For instance, when more general overall information was collected through questionnaires and the Delphi technique, all schools were asked to participate, but during the more in-depth examination of the use of data in the schools through observations, interviews and reflective journals, a sub-sample of schools were selected. The specific sampling for each cycle is discussed in Chapters 5-8 as applicable to the cycles focused on in the respective chapters. The overall research procedures are introduced in the next section.

4.2.5 Research Procedures

Figure 4.5 (based on the representational style of Mafumiko, 2006) details the design and evaluation interactions for this research. The feedback prototypes are illustrated in blue and the evaluation activities in green. Every full design cycle consists of both the prototype development and adaptation followed by the implementation[28] and the corresponding formative evaluation of that prototype. The research procedures for each cycle are discussed separately with the data for each cycle in Chapters 5-8.

4.2.6 Shifts in Emphasis in the Design Research Process

The shift in emphasis quality criteria during the cycles are illustrated in Table 4.2 by way of shading (based on the representational style of Nieveen, 2009, p. 96). The table also illustrates the instruments that were employed to address the research questions[29] and evaluative criteria.

Over and above fulfilling evaluative criteria, design research must also fulfil the required methodological norms to ensure a rigorous study. The next section discusses the methodological norms and how they were applied in this study.

[28] Some design research authors such as Plomp (2009) prefer the term 'try-out', while others such as Reeves, et al., (2005) prefer testing the solution in context. In this thesis the term implementation is used as the feedback system was functioning during the entire design process and had to provide the feedback on the performance monitoring data to all schools, twice a year.

[29] Research sub-question 2: *How can an existing learner performance monitoring system be appropriately adapted, contextualised and translated to the South African context?*" was addressed in full in Chapter 2 prior to the commencement of the design research process. It is therefore not represented in Table 4.2.

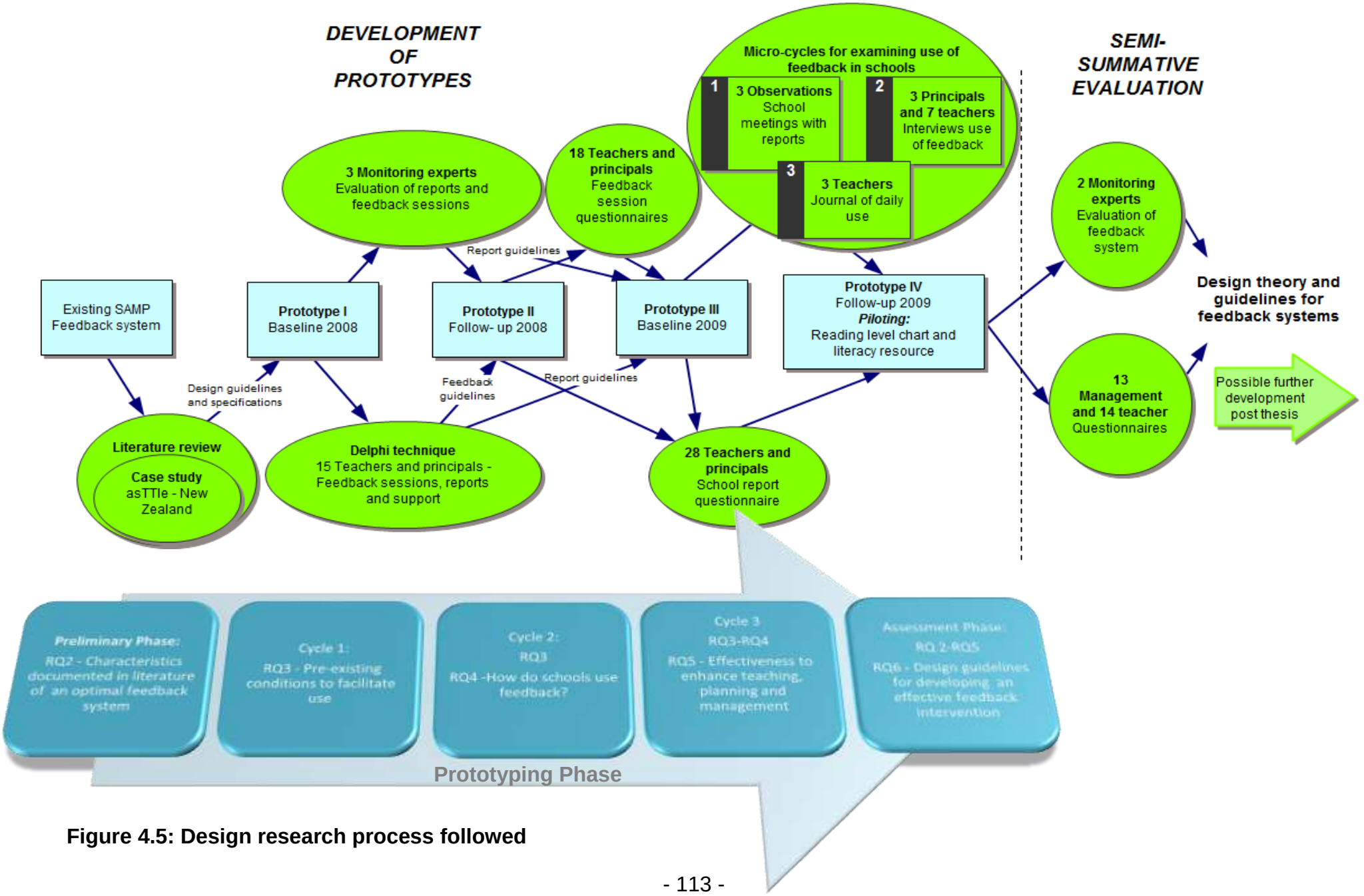


Figure 4.5: Design research process followed

Table 4.2: Quality emphasis per development stage

QUALITY CRITERION	DESIGN SPECIFICATIONS AND GLOBAL DESIGN OF FEEDBACK SYSTEM	ESTABLISHING CONDITIONS FOR USE	TRANSFORMING CONDITIONS FOR USE INTO ACTION	COMPLETE INTERVENTION
Research question focus	***Research question 2:*** *What are the characteristics documented in literature of an optimal feedback system for use in school-based monitoring?*	***Research question 3:*** *What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?*	***Research question 3 (limited report questionnaire)*** ***Research question 4:*** *How do schools use feedback?* ***Research question 5 (focussing on expected efficacy):*** *How effective is the feedback system enhancing classroom practices, management and planning activities?*	***Research questions 3-5***
Evaluation Cycle	***Formative – Cycle 1***	***Formative - Cycle 2*** ***Formative - Cycle 3*** ***Formative – Cycle 4***		***Semi-summative – Cycle 5***
Relevance	× Exemplary Case study × Literature review	× Questionnaires × Delphi × Expert appraisal	× Participant observations – planning meetings × Semi-structured reflective journals × Semi- structured interviews	× Questionnaires × Expert appraisal
Consistency	× Exemplary Case study × Literature review	× Questionnaires × Delphi × Expert appraisal	× Participant observations – planning meetings × Structured reflective journals × Semi- structured interviews	× Questionnaires × Expert appraisal
Practicality: ***Expected***	× Exemplary Case study × Literature review	× Questionnaires × Delphi × Expert appraisal	× Participant observations – planning meetings × Semi-structured reflective journals × Semi- structured interviews	× Questionnaires × Expert appraisal
Actual		× Questionnaires	× Participant observations – planning meetings × Semi-structured reflective journals × Semi- structured interviews	
Effectiveness: ***Expected***	× Exemplary Case study × Literature review	× Questionnaires × Expert appraisal	× Participant observations – planning meetings × Semi- structured interviews	× Questionnaires × Expert appraisal
Actual			× Participant observations – planning meetings × Semi- structured interviews	
RESEARCH QUESTION 6: ***Which design guidelines for the development of an effective feedback intervention for school-based monitoring can be identified?***				

4.3 Methodological Quality

As a pragmatist, the choice of paradigm is viewed as secondary to the usefulness of the methods employed to address the research questions appropriately (Mertens, 1998; Onwuegbuzie & Johnson, 2004; Tashakkori & Teddlie, 1998). At the same time, pragmatists acknowledge that it is inappropriate to evaluate quantitative methodologies according to qualitative norms and vice versa. In this study the norms specific to either quantitative or qualitative methodologies were employed, dependent on which methodologies were applied during any given phase.

The quantitative methods were subject to the norms of validity, reliability and generalisability. Traditionally, validity has been categorised as either content validity, criterion validity or construct validity. This separation has laid the foundation for a common misconception that establishing only one type of validity is sufficient. Messick (1989) argues that validity should be seen as a unitary concept that is established through several sources of evidence including content and criterion validity. Messick (1989) further argues that validity should go beyond these types of investigations to include the social context in which the examinations take place. Therefore, the validity issues should extend to include an examination of not only the evidential basis for a particular interpretation of the data, but also the consequences of a particular interpretation. This unitary concept of construct validity therefore includes relevance or utility, value implications and social consequences (Messick, 1989). This view of validity is appropriate for this study into optimising a feedback system for appropriate use and action.

It was this extended, unitary conceptualisation of validity that was employed with the quantitative components of this research. The instruments were designed to measure the underlying constructs of relevance, consistency, practicality and effectiveness of representation in reports, feedback sessions and support provided to schools. During the development of the questionnaires, items were mapped according to the variables that each question would address. The content and face validity of the instruments were evaluated by research psychologists and monitoring experts. In this case, the

consequential validity relates to how the data from the instruments were used or effectiveness criteria. Data from the instruments were employed to improve the quality of the feedback system aimed at improving the way education is provided in the schools. The quantitative component of this study therefore fulfils the methodological norms for the quantitative methods.

In terms of the qualitative components of this research, the norm of trustworthiness as first suggested by Guba and Lincoln (1985, pp. 289-331) was adhered to. Trustworthiness is established by ensuring transferability, credibility, dependability and confirmability as depicted in Table 4.3. This approach relies on 'thick'[30] descriptions to allow other researchers to transfer results to their own context. Prolonged engagement, referential adequacy, peer debriefing and member checks are employed to establish the credibility of the research. Finally, an inquiry audit trail is made accessible to readers and peers in order to examine the dependability, credibility and confirmability of the data.

In this research, the role of the researchers was dual, as the researcher acted both as the monitoring and feedback facilitator and the evaluator as design researcher. The implications of this dual role for the methodological quality of the study are discussed in the next section.

[30] The term thick description originates from the work of anthropologist Clifford Geertz (1975) and means that the description of behaviour should include the context, to make the behaviour meaningful to an outsider

Table 4.3: Establishment of trustworthiness in qualitative research

CONSTRUCT	ATTAINMENT IN THIS STUDY	EXPLANATION OF THE SPECIFIC APPLICATION IN THIS RESEARCH
Transfer-ability	■ Thick descriptions	■ The *Atlas.ti* qualitative data analysis programme was used to provide outputs in the form of rich descriptions that are contextually situated (Scientific Software Development, 1997).
Credibility	■ Prolonged engagement	■ The project was first initiated in 2003. The researcher has been involved with the original SAMP project and the schools in the project since 2006. This particular PhD study incorporated another 36 months of data collection alone.
	■ Referential adequacy	■ Various data sources were used to ensure that the data are representative. This approach is also referred to as triangulation (Mertens, 1998; Willig, 2001). In this case, triangulation was achieved through the incorporation of interviews, observations, reflective journals, the Delphi technique and questionnaires administered to principals, teachers and experts. There is thus not just a triangulation of data sources, instruments and theories, but also of method.
	■ Peer debriefing	■ Supervisors and various colleagues played the role of 'devil's advocate' questioning the research process and methodology continuously.
	■ Member checks	■ Member checks took place throughout the design generation and research text production. Member checks extended beyond just validating transcriptions, to discussions of interpretation particularly during the evaluations of design cycle 4 of Prototype III.
Dependability and confirmability[31]	Guba and Lincoln (1985) suggested the concept of inquiry audit to establish both dependability and confirmability. Inquiry audit refers to a process where an auditor examines documentation in order to attest that the findings are supported by the data and are internally coherent.	Supervisors and critical readers were approached to examine the following data in this research: ■ Raw data: Interviews, tapes and field notes. ■ Data reconstruction and synthesis products: In the form of the *Atlas.ti* outputs, in the form of a web pages including code networks. ■ Process notes: these notes took the form of memos generated by me during the data analysis as well as sections of my personal research diary incorporated into *Atlas.ti*. ■ Material relating to intentions and dispositions: Personal notes and expectations were represented in the *Atlas.ti* memos. The audit trail will be made available with the thesis on the audit trail DVD.

(Babbie & Mouton, 2001)

[31] Refers to whether it is possible to investigate the data, analysis and synthesis products of the research to establish whether the interpretations of the researcher are justifiable and consistent.

4.3.1 Role of the Researcher

Riessman (1993, p. v) states, “the construction of any work always bears the mark of the person who created it”. The mere choice of inclusion or omission of aspects in the study is part of the researcher signature. Geertz (in Clandinin & Connelly, 1994, p. 424) warns “[t]oo vivid a signature runs the risk of obscuring the field and its participants; too subtle a signature runs the risk of the deception that the research text speaks from the point of view of the participants.”

As the researcher, I have multiple roles in this research, explored in this section along with an explanation of the role of self-reflexivity in managing these roles and the tensions that may arise from them. This section gives the reader an idea of the ‘lenses’ I wore during the research and the influence they may have had on the process.

My role in this research process was multi-fold as:

- A full-time researcher at the CEA charged with the project coordination for the SAMP project. While engaged in this doctorate, I was employed to ensure the smooth running of the SAMP project. This was essential to produce processes and data that the feedback system employs. I was therefore not only the researcher, but also the facilitator of the monitoring and feedback process.
- An educational psychologist who has worked in schools, but not as a teacher.
- The person responsible for the development and evaluation of the various feedback prototypes for this thesis.

As I was responsible for the coordination of SAMP, the development and evaluation of the prototypes and the implementation of the prototype, there may have been some conflict of interest. Co-researchers such as teachers, principals, HoDs and experts in the field of monitoring and utilisation were enlisted to support the evaluations and guard against bias. Plomp (2009) suggests several other methods to compensate for possible conflict of interest in the design research process that was also employed in this study:

- opening the research to professional scrutiny to people outside the project

- shifting from the 'creative designer' perspective with a focus on creativity and improvement of the feedback system, to the 'critical researcher' perspective in order to critically evaluate the success and value of the system as the research progresses

- having a good quality of research design, e.g.

  - *strong chain of reasoning* (Krathwohl, 1998) - each part of the research design has equal importance

  - *triangulation* – of methods (qualitative and quantitative), data sources (principals, HoDs, teachers and evaluators and investigators not only during data generation, but also data analysis and interpretation

  - *empirical exploration* of both the usability and the effectiveness of the intervention

  - *systematic documentation, analysis* and *reflection* of the design, development, evaluation and implementation process and their results

  - *application of a variety of methods and tactics:* e.g. use practitioners and other researchers as 'critical friends'; use multiple observers or raters and calculate inter-observer or rater reliability.

  - *ensuring validity and reliability of data and instruments*

As researcher, I also have the power of presentation through my involvement in the project and the authorship of this thesis. Geertz (in Czarniawska, 2004, p. 108) poignantly describes the burden of authorship: "[it] cannot be evaded, however heavy it may have grown; there is no possibility of displacing it onto 'method', 'language', or… 'the people themselves' redescribed … as co-authors". This inevitable burden of authorship cannot be denied, but needs to

be acknowledged and discussed to inform the reader of my position and approach to this responsibility.

My worldview and the research relationship influence the representation in the research text (Willig, 2001, p. 53). Bullough (in Goodson & Sikes, 2001, p. 56) explains that although interpretations will inevitably vary "we cannot write just anything we wish, … interpretations, however tentative must be disciplined by data, and … we must proceed cautiously and carefully". In order to make it possible for the reader to investigate the basis for my interpretations, the research audit trail is enclosed on a DVD in this thesis.

Research diary entries are used throughout the text to show my own positioning through self-reflexivity. Gouldner (in Plummer, 2001, p. 207) suggests that reflexivity involves analysis of one's own work and reflection through the same tools one uses to examine participants. This approach to self-reflexivity is used by me in this research and is shown through diary entries and references to analysis memos created in *Atlas.ti*. Such notes are reflected with an in-text citation e.g. (Research diary: 01/09/2007) as already seen.

This research aims to have implications for learner performance monitoring feedback systems in the Tshwane region, as well as for South Africa, and possibly internationally. In the next section, the realm of application of the research findings for this study is discussed.

4.3.2 Realm of Application

This research is conducted with the aim of finding application and influence beyond just the current scope of the study. In quantitative methodologies this is achieved through statistical generalisability. However, generalisation to a population can only be achieved if the sample is representative of the population and is large enough to capture the qualities of the population to which the results are to be generalised. In the case of this study, the population were schools in the Tshwane area teaching in the three languages included in the SAMP project. As discussed previously (see Section 4.2.4), the sample was stratified and although representativeness was sought, the

sample was relatively small and representativeness cannot be fully substantiated. It is therefore not possible or necessary to achieve generalisability in this study.

In qualitative research transferability is usually sought by providing thick descriptions of the research process, the findings and context, thus allowing a reader in another context to determine if the study is applicable and useful in his or her own context. This research however represents a mixed methods study that acknowledges the importance of context in studying and optimising the use of any feedback system. This recognition has implications for generalisability, as Cronbach (1975, p. 125) states: “When we give proper weight to local conditions, any generalization is a working hypothesis and not a conclusion”. Therefore, the concept of analytical generalisability as opposed to statistical generalisability and transferability was sought.

Analytical generalisability relies on thick description of the research, findings and context to allow other researchers to replicate the study in different contexts (Yin, 2003). The design principles were thus tested in a variety of contexts or cases through replication. Analytical generalisability thus means that if the design principles can be accepted for a number of cases, or contexts, those principles might be accepted for a much larger number of contexts. In the case of this study design guidelines were provided along with rich descriptions of the research to allow other researchers to apply the research to their own feedback system contexts. When the research is transferred and repeated in a number of different contexts, the research will achieve higher levels of analytical generalisability.

4.4 Conclusion

In this chapter the research design for this study is introduced with specific emphasis on the design research methodology. The choice of the pragmatist paradigm was discussed and the shifting focus on the design and evaluative criteria explored, along with the different roles of the research collaborators in the design and evaluation process. The methodological quality of this study was also elaborated upon, with a discussion of the dual role of the researcher

as facilitator and evaluator and the realm of application of the research. Chapters 5-8 will focus on detailed discussion of the research procedures results, findings and design guidelines for each cycle.

CHAPTER FIVE
Preliminary Phase: Problem identification, needs and context analysis

> The case study is the method of choice when the phenomenon under study is not readily distinguishable from the context. The inclusion of context as a major part of a study… creates distinctive technical challenges. First, the richness of the context means that the ensuing study will likely have more variables than data points. Second, the richness means that the study … will likely rely on multiple sources of evidence. Third… distinctive strategies will be needed for design and analysis. (Yin, 2003, p. 5)

This chapter takes a closer look at the research design and results of the first of five design cycles for this thesis. The first cycle is known as the Preliminary Phase and addresses sub-question 3: "What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?" Each design research cycle presents a small, but complete research process that allows the identification of design principles that inform the next cycle of development. The initial cycle is described as the Preliminary Phase (Plomp, 2009) of the design research process. This phase focused on conceptualising the feedback system and defining the design specifications The information derived from the literature review (Chapter 3) as well as the *Prior Development, Needs and Context Analysis* and *Exemplary Case Study-asTTle, New Zealand* were combined and are discussed in full in this chapter.

The Preliminary Phase is discussed with reference to the specific research questions and evaluative foci, the pre-existing system being evaluated, the research design, results and design principles from the phase. The focus of this chapter in terms of the overall research cycles is discussed in Section 5.1. Section 5.2 introduces the pre-existing feedback system (Sections 5.2.1-5.2.3) and the evaluation of the pre-existing system is discussed in Section 5.2.4. The design principles derived from the literature in Chapter 3 are provided in Section 5.2.5, followed by the exemplary case study of the asTTle system in

NZ (Section 5.2.6). Section 5.2.7 introduces the design principles derived from the case study.

5.1 Research Cycles

This thesis consists of five full design cycles. The *Preliminary Phase* starts with the existing feedback system and consists of one design cycle. The *Prototyping Phase* consists of three full design cycles that are discussed in Chapter 6 and 7, while the *Assessment Phase* consists of one cycle represented in Chapter 8. The design and evaluation interactions for this research are illustrated in detail in Figure 5.1. The pre-existing system and feedback prototypes are shown in blue and the evaluation activities in green. Every full design cycle consists of the prototype adaptation followed by implementation and corresponding formative evaluation of that prototype.

Each cycle represents a complete, small research and is therefore discussed separately (Chapters 5-8), beginning with a description of the prototype, followed by a closer examination of the research design used to evaluate the prototype. The evaluation activities are guided by the research questions addressed and the evaluative focus for each specific cycle. While there is a specific focus for each cycle there is always some overlap and a cycle may address some aspects of other research questions or evaluative foci. The research methods for each cycle include the sampling (Section 5.2.6.2), data collection and capturing (Section 5.2.6.3-5.2.6.4), analysis (Section 5.2.6.5) and discussion (Section 5.2.6.6).

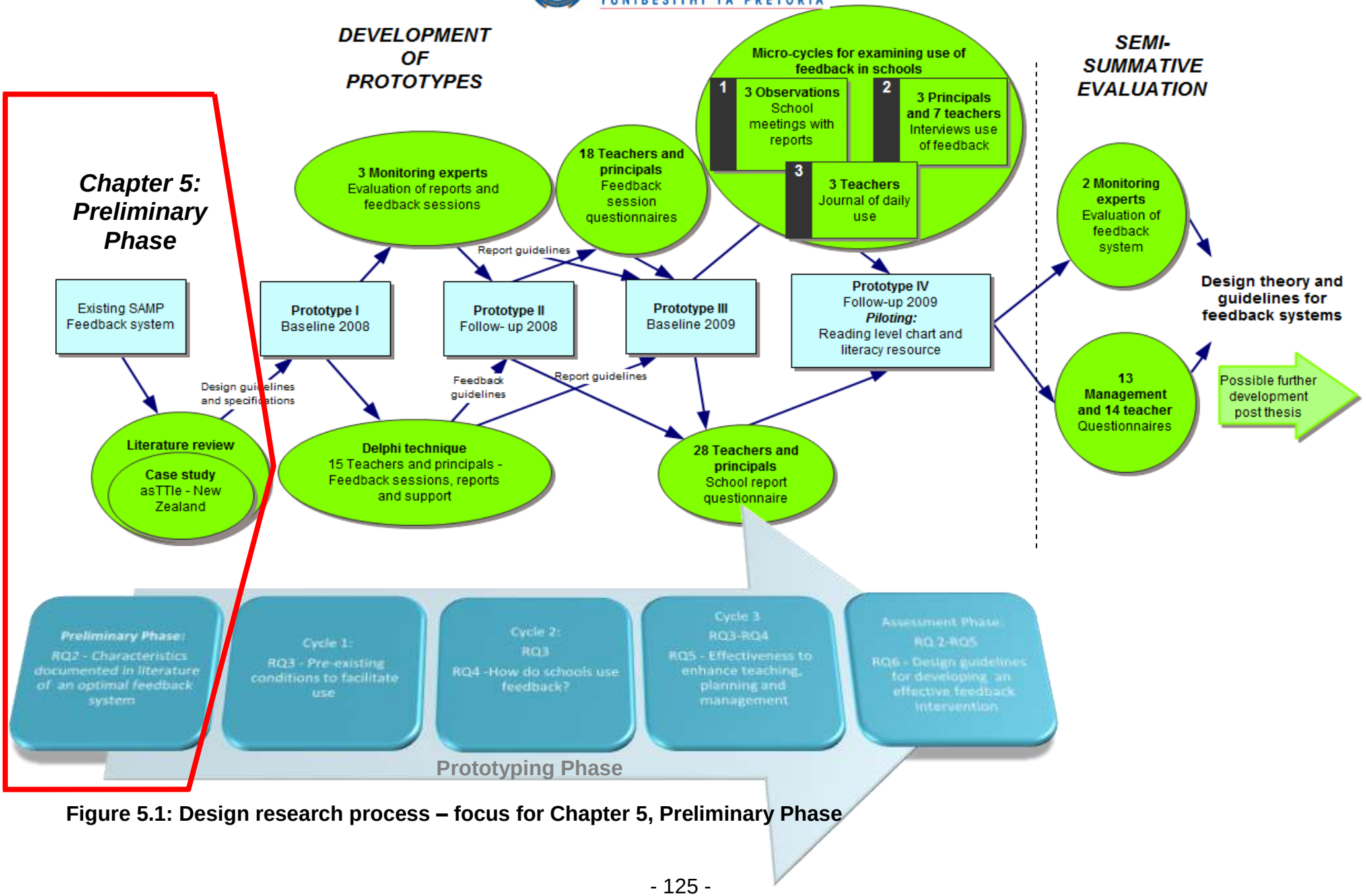


Figure 5.1: Design research process – focus for Chapter 5, Preliminary Phase

5.2 Prior Development, Needs and Context Analysis

Schools participating in the existing feedback system prior to 2006 expressed a need to receive feedback more quickly and that the data be presented in such a manner that it is easier to use for planning, decision-making and action in the school environment. It was necessary to develop design guidelines and criteria for the global design of the feedback system to facilitate use of the data. In order to accomplish this a number of approaches were employed including a literature review of selected international SPFSs (see Chapter 3).

This was supplemented with an exploration of an exemplary case in the form of the asTTle (assessment Tools for Teaching and learning) system in NZ. By studying the feedback system used with asTTle in New Zealand, thus gaining a better understanding of how context influences the effective use of a feedback system. The Preliminary Phase therefore pertained to the second[32] ***research question***:

2. *What are the characteristics documented in literature of an optimal feedback system for use in school-based monitoring?*

Design research is a parallel process with an intervention or product being designed along with design principles to further the body of knowledge. The design principles are generated out of the evaluation of the prototype or for this first cycle, the pre-existing system. However, prior to the evaluation of the cycle the researcher has a conceptualisation of the class or type of design guidelines that are sought in the cycle. The evaluation information for the first cycle was therefore aimed at generating ***design guidelines*** relating to the design specifications for the global design of the feedback system:

1. *Design specifications and global design:* Determining the components necessary for the optimised feedback system. Generating a preliminary conceptual framework for the feedback system.

[32] Research question one: "How can an existing learner performance monitoring system be appropriately adapted, contextualised and translated to the South African context?" was addressed in full in Chapter 2.

The evaluation in this cycle focused specifically on the ***evaluative foci*** (as discussed in Section 4.2.2) of relevance and consistency:

1. *Relevance (content validity):* The system and its design should be based on state-of-the-art (scientific) knowledge. The feedback system must clearly be connected, in form and purpose, to the learning performance monitoring system for which it provides feedback.
2. *Consistency (construct validity):* The system must be 'logically' well designed. The various parts are well defined and the connections between them explicitly postulated. Although some elements may emerge more clearly throughout the process, there are no internal contradictions.

In the following section, the pre-existing feedback system that formed part of PIPS (Performance Indicators in Primary Schools) is introduced briefly, along with the design principles from the informal evaluative activities of this cycle. This is followed by the guidelines derived from the literature discussed in full in Chapter 4. Then the research procedures and data for this cycle are discussed.

5.2.1 Pre-Existing Feedback System (Prior to 2006)

Prior to 2006, the data for the feedback system were generated through a computerised version of PIPS that was translated into Afrikaans and Sepedi. Learners from seven schools were assessed in English, Afrikaans or Sepedi. Fieldworkers were provided with hired laptops on which the assessment software was loaded to conduct the fieldwork. Around 47 learners were assessed per school resulting in 327 learners assessed in 2005. Each learner was assessed on a one-to-one basis by a fieldworker who used the computerised testing programme. Data were downloaded from the laptops and the code sent to the CEM in Durham.

Once the data were processed in Durham, a dataset was sent back to the CEA for further analysis and feedback. The system still contained some programming errors and bugs, which meant that for the translated languages such as Afrikaans and Sepedi as well as the South African English, not all the data from the assessment could be used. In all three languages part of the Letters and

Vocabulary subtests had to be omitted from the calculations as the programme terminated prematurely, while the reading items in Afrikaans still provided difficulty along with the more advanced Sepedi reading items (Eiselen, 2005a). This meant that only the mathematics scores were comparable across the different languages. The project leader would produce a report for each participating school based on the usable data and once the reports were completed, a feedback session would be arranged at the University of Pretoria. In a given year a school participating in the PIPS system would have its Grade 1 learners assessed at the beginning of the year and based on this a report for their school would be produced. The principals would then be invited to a feedback session at the University of Pretoria where the reports were handed out. This process was repeated again at the end of the year for the follow-up assessment. The components of the Pre-existing PIPS feedback system are shown in Table 5.1.

Table 5.1: Pre-existing PIPS feedback system components

Component	Description
Paper-based baseline report	Produced for each school individually and handed out at the feedback session
Baseline feedback session	Principals from all participating schools invited to the University of Pretoria
Paper-based baseline report	Produced for each school individually and handed out at the feedback session
Follow-up feedback session	Principals from all participating schools invited to the University of Pretoria

5.2.2 Reports – Pre-Existing Feedback System

The report outline for the pre-existing feedback system was similar for the baseline and follow-up reports:

1 Introduction
2 Introduction to the Assessment
  2.1 Early Phonics
  2.2 Early Reading
  2.3 Early Maths
3 Interpretation of Results
4 Overall Trends Observed
  4.1 Overall Achievement in Maths
  4.2 Overall Achievement of xxx School
  4.3 Achievement at Different Ages
  4.4 Achievement Boys and Girls
  4.5 Which Items Did Learners Find Difficult and Which Were Easy
    4.5.1 Phonics
    4.5.2 Reading
    4.5.3 Maths
5 Individual Learner Results
  5.1 Learners at Risk and Exemplary Learners For Each Category
    5.1.1 Phonics
    5.1.2 Reading
    5.1.3 Maths
6 Conclusion
7 References

Figure 5.2: Report outline – pre-existing system

Both the baseline and follow-up reports of the pre-existing feedback system (demonstration reports can be found on the audit trail DVD) started with a short description of the project, followed by a description of the various subtests (see Figure 5.3).

Vocabulary

Vocabulary is assessed through two picture items. The learners are shown the picture, such as the one in Figure 2, and asked by a voice on the computer, to point to common objects in the picture. For example in Figure 2 the learner is asked to point to some carrots. If the learner is able to complete this task the field worker will record this response as "right".

Figure 2: Vocabulary subtest:
The learner is asked to point to a picture on the screen

Figure 5.3: Description of the Vocabulary subtest – pre-existing system

All results were produced as category scores with 1 = 0-25%; 2=26-50%; 3=51-75%; 76-100%. First, the Mathematics results across the different schools were shown (see Figure 5.4), these being the only set of comparative results provided, as the sample was too small to compare the different scale results per language group and the assessment were not shown to be equivalent across the other scales.

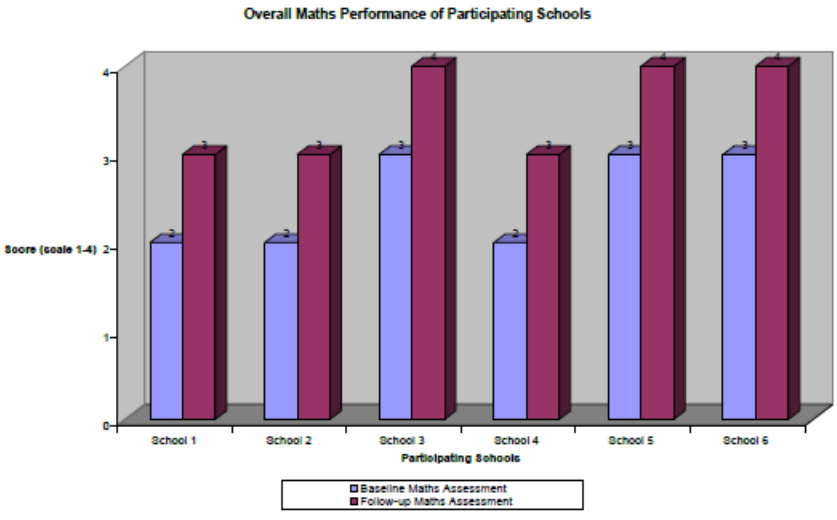


Figure 19: Overall Maths Achievement of all the schools which participated in this years study (2005)

XXX School is indicated as school 6 on the figure above. Although school XXX achieves of the highest mathematics scores, it also receives learners, at the beginning of the year, whose early mathematics skills are already partially developed.

Figure 5.4: Mathematics Scale results across schools - pre-existing system

This was followed by the school's scale scores represented in a bar graph format. The graph showed the school's comparative performance separately for Early Phonics, Early Reading and Mathematics scale with the results of both the baseline and follow-up assessment depicted (see Figure 5.5). Bar graphs depicting the comparative performances for each scale according to gender and age for the school were also provided.

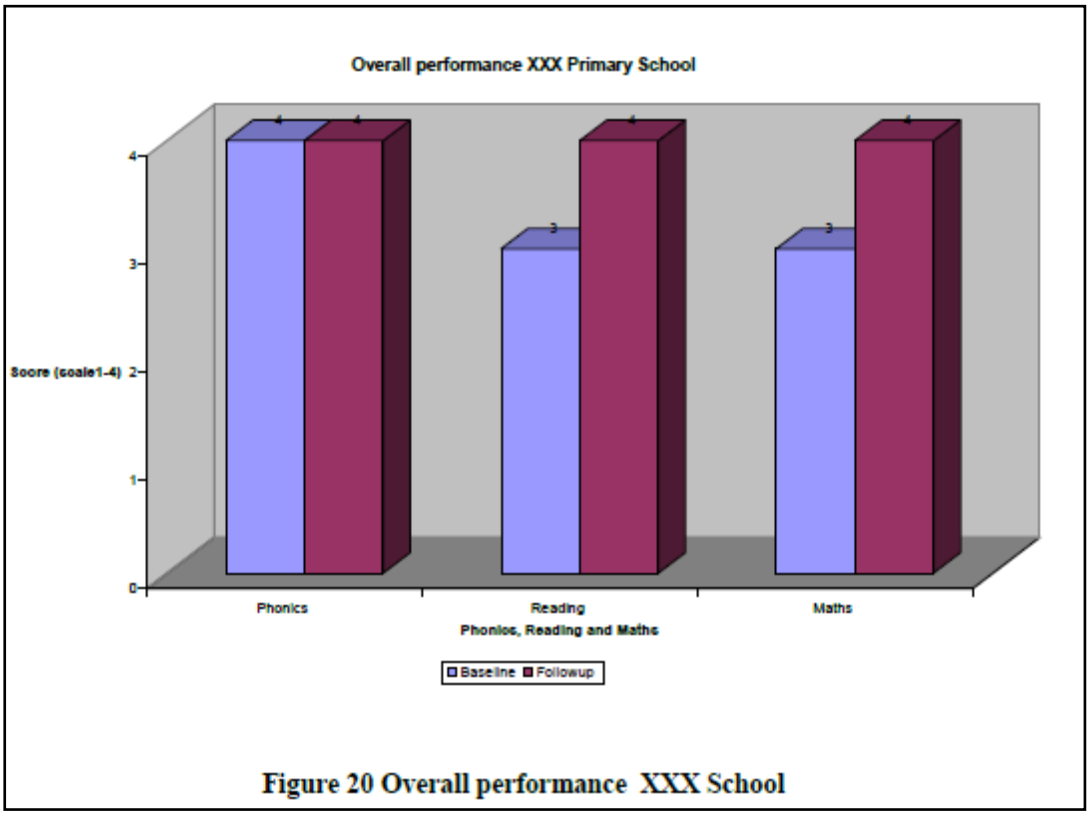


Figure 5.5: Overall performance for the school - pre-existing system

The bar graph representations were followed by a tabulation of the easiest and most difficult subtest item for each school (see Figure 5.6). In total, six such tables were produced, one for the easiest and one for the most difficult items for each of the three scales. This concluded the reporting of overall school results.

Figure 24: The easiest reading items

Subtest	The easiest item of this subtest for all the learners who participated in this years study	The easiest item of this subtest for XXX School learners
Vocabulary	"Can you point to a fork?"	"Can you point to some carrots?"
Letters	"What is this letter?" S	"What is this letter?" S O
Mix up	"Which word says rabbit?"	"Which word says rabbit?"
Quiz words	"Which word says car"	"Which word says car"

Figure 5.6: Easiest Reading easiest items – pre-existing system

Thereafter there was a section on individual learner level, which started with a tabular representation of learner results presented in categories from 1-4 for each of the three scales for the baseline and follow-up results (see Figure 5.7). The table was followed by a section in which learners were listed as having trouble or achieving exceptional results in each of the three scales.

Table 3 Performance of Mrs XXX's class

Learner name	Learner surname	Baseline Phonics	Follow-up Phonics	Baseline Reading	Follow-up Reading	Baseline Maths	Follow-up Maths
ZZZZ	YYY	3	1	1		2	3
ZZZZ	YYY	2	3	1	1	2	2
ZZZZ	YYY	2	4	1	1	3	2
ZZZZ	YYY	3	3	1	2	3	3
ZZZZ	YYY	3	2	2	1	3	3
ZZZZ	YYY	3	3	1	2	3	3
ZZZZ	YYY	3	3	1	2	3	4

Figure 5.7: Individual learner result tables – pre-existing system

Once the reports were completed, a feedback session was arranged at the University of Pretoria.

5.2.3 Feedback Sessions – Pre-Existing Feedback System

Schools were invited to attend and received the reports at the feedback session, which took about two hours and consisted of a presentation by the project leader pertaining to:

- Introducing the CEA and the project
- Describing the assessment

- Presenting the aggregated results across all the schools
- Indicating the most difficult and easy items on average for the schools
- Presenting the future plans for the project
- Providing an opportunity for questions

The session was concluded with light refreshments and open discussions between the schools and research team.

5.2.4 Informal Evaluation of - Pre-Existing Feedback System

Two schools indicated that they wished to cease participation in the PIPS project. Informal face-to-face and telephonic conversations were conducted with these schools and the other five schools. Schools raised a number of concerns about the reports and feedback sessions:

5.2.4.1 Reports

Schools indicated that they found the PIPS system useful in a number of respects:

1. Teachers indicated that the individual learner results were very useful. These results were used for comparison with their own standard of marking.
2. The identification of learners who had trouble with certain sections of the assessment or who achieved exceptionally well was also useful. Teachers employed this information to determine which learners needed individualised support or more advanced stimulation.

At the same time, the schools expressed a number of concerns:

1. The only comparative data reported on between schools were on the Mathematics Scale. Due to the errors and bugs in the computerised system and because the equivalence across languages had not been established, the Reading and Phonics scores could not be compared across the different languages. The small sample size also meant that comparative data for schools of the same language group (maximum of three per language) could not be provided as this would impinge on

confidentiality. Therefore, only the Mathematics data were presented in comparison to other schools in the assessment. This meant that schools could not determine how their schools were doing overall or on the Phonics or Early Reading scales compared to other schools. This made it difficult for schools to decide what areas they should attend to in their curriculum planning.

2. The use of the large categories (25% range) in the school scores also meant that often no growth would be seen in the overall school graphs. See for example the phonics results shown in Figure 5.5 (above).

3. Once a particular area such as Early Reading was identified as a concern no data were available to determine which subtests were of the greatest concern. Therefore, the school still did not know which particular aspects of reading should enjoy greater attention.

4. Schools also indicated that while the identification of the easiest and most difficult items in each subtest was interesting, it had little impact on their planning and classroom practice.

5. Teachers and principals with African learners in their schools expressed doubt about some of the results as the graphics used in the computer programme and illustrated in the reports were considered by them as Eurocentric and foreign to some of the learners.

6. The results of individual learners were difficult to relate to planning for the class overall.

7. The lack of a recommendation section made it difficult for schools to gain an overall view of their performance and to take action.

8. Schools also expressed concern about the long time span between the assessment and reporting. As the data went through a long process of cleaning and analysis involving both the CEM and the CEA, report writing was delayed. Reports were also manually produced per school; which meant that reports were often only received three months after the assessments were concluded.

9. Two of the schools indicated that there was a breakdown in communication between themselves and the previous project coordinator.

5.2.4.2 Feedback Sessions

The schools expressed that it was a novelty to have contact with the University and hear about the project. They however experienced the sessions as very long and repetitive as the focus was mainly on the format of the assessment and the different subtest. While the feedback sessions helped with understanding the data, they did not support them with transforming the data into action. The need was for feedback sessions that supported school improvement action based on the data.

5.2.5 Design Principles from Literature Review

Based on this initial analysis of the pre-existing feedback system, a literature review was conducted to determine tentative design principles for the optimisation of the pre-existing feedback system (Discussed in full in Chapter 3). As no such system existed in South Africa at the time of the optimisation, international SPFSs (School Performance Feedback Systems) which had been documented over an extended period of time were reviewed. The four cases included the CEM Suite (UK), the asTTle system (NZ), Zebo (Netherlands) and SAM (Louisiana).

Although these countries have varying approaches to SPFSs, there are some principles that emerge (the principals are mapped to the systems in Table 5.2).

1. The data must ***not be viewed as part of unfair high-stakes accountability practises***. Its main purpose should be as a driver for improvement and not accountability (Hattie, 2005; Tymms & Albone, 2002).

2. ***School and educator expertise should be utilised in the development*** and improvement of the feedback system to ensure contextual appropriateness and a sense of ownership (Hendriks, et al., 2001; Tymms & Coe, 2003).

3. The feedback system must ***provide tools to support school improvement-driven practices*** and support greater school autonomy (Angelle, 2004; Hendriks, et al., 2001; Teddlie, et al., 2002).

4. Data must allow for ***comparison of a school's performance to other groups*** (Angelle, 2004; Crooks, 2002; Hattie, 2005; Hendriks, et al., 2001; Tymms & Coe, 2003).

5. A ***short turn-around time*** from assessment to reporting is essential to ensure the data is still relevant.

    - ***Use of ICT is important to improve turnaround time*** and increase schools sense of autonomy (Angelle, 2004; Hattie, 2005; Hendriks, et al., 2001; Teddlie, et al., 2002).

Table 5.2: Principles emerging from the international SPFSs

Principle	UK - CEM Suite	NZ – asTTle	Netherlands - Zebo	Louisiana – SAM
Not part of high-stakes accountability	✓	✓	✓	
School expertise used in development	✓	✓	✓	✓
Provide tools for improvement practice	✓	✓	✓	✓
Comparison with other groups	✓	✓	✓	✓
Importance of short turnaround time	✓	✓	✓	✓

The literature on SPFSs discussed in Chapter 3 provided a rich source from which to develop initial design guidelines for the components necessary for the feedback intervention. However, investigating contextualised processes and logistical issues through literature alone was insufficient. It became clear that contextualised knowledge about the functioning of the SPFS in its particular context was required in order to transfer the learning to the South African context. A case study of an exemplary feedback system in its context was conducted to gain this knowledge.

5.2.6 Exemplary Case Study – asTTle, New Zealand

A specific feedback system and context had to be identified for the exemplary case study. The focus of the case study was to expand on the understanding of

the components and aspects of a good feedback system and its functioning, in context holistically within one education system to further development of the design guidelines. Limited resources meant that such a case study could only pertain to one system. It was therefore essential to identify the exemplary case to study selectively.

An examination of PIPS in England may have been appropriate for such a study. This option was rejected as regular communication already takes place with the CEM in Durham about the Value-added project. Also, much information had already been gathered by two previous researchers on the project during a three-week study tour. The communication has already had an influence on the development of the feedback system without a specific study of the UK case. Visits by academics between the CEM and CEA also take place as part of the Value-added project. Likewise, knowledge of the Dutch ZEBO system had also already been accessed through the Dutch consultants involved in the SANPAD collaboration and previous visits by colleagues to the Netherlands. While SAM functions well in Louisiana, it only covers one state and not the entire educational system. SAM also forms part of high-stakes accountability practices which were avoided in the pre-existing system. The asTTle system thus presented an opportunity to explore another monitoring and feedback system that had not yet had a direct impact on the development of the feedback system.

The reasons for selecting asTTle for the exemplary case study are set out below:

- asTTle has been successfully employed in New Zealand since 2001
- asTTle has been lauded as a success by both the New Zealand Ministry of Education and New Zealand schools
- asTTle is widely used in schools in New Zealand
- New Zealand utilises Outcomes Based Education (OBE), as South Africa does
- The New Zealand education system is relatively small, making it more feasible to access users and stakeholders at various levels in the system within a short period. This makes it possible to gain a holistic picture of a national system.

- The DoE in South Africa was also exploring the possibility of adapting asTTle in some form

  (Crooks, 2002; Hattie, et al., 2004b; New Zealand Ministry of Education, 2007a)

The asTTle system was therefore selected and a month-long visit to NZ undertaken to study the use of asTTle in context. In the next section, the literature review for the case study is discussed, followed by the research procedures with reference to sampling, data collection, data capturing, data analysis and a discussion of the data.

5.2.6.1 Literature review

The South African and New Zealand education systems both had similar outcomes-based education curricula at the time and both share policies that encourage assessment for learning as opposed to assessment of learning (Gardner, 2006). In other respects, there are vast differences between the two countries. For example, South Africa is still a developing country, with grave internal infrastructural and economic disparities, which complicate the effective implementation of education policies. Another significant difference is that academic performance in New Zealand, by international measures of mathematics and science such as TIMMS 1995, 1999 and 2003, as well as reading, PIRLS 2006, is considerably higher than that of South Africa (Harmon, et al., 1997; Martin M. O., et al, 2000, 2004; Mullis, et al., 2000, 2004). The South African Educational context is discussed in depth in Chapter 3, and the NZ educational context here.

New Zealand educational context

NZ participates in national and international assessments in education as part of its system monitoring agenda. NZ generally performs exceptionally well on international assessments, for example with performance in PIRLS 2006 (Chamberlain, 2007), PISA 2003 (New Zealand Ministry of Education, 2007b) and TIMSS 2003 (New Zealand Ministry of Education, n.d.) having been well above average, and consistently so (New Zealand Ministry of Education, n.d.). While the NZ education infrastructure is far superior to that of South Africa, there

are interesting differences in how the nation has implemented the assessment for learning policy.

The NZ education system has undergone several changes in the last two decades, the most significant being a shift from centrally controlled education through a department of education to autonomous schools. This was accomplished in 1989 through the Tomorrow's Schools reform that transferred operational control of schools to the boards of trustees of individual schools (Crooks, 2002; Fiske & Ladd, 2001). Under the Ministry of Education (MoE), government funding of schools continues, within a framework of self-governance, self-management, and self-directed improvement. The decentralisation of education increased fears for the maintenance of national standards, especially since NZ's outcomes-based education system had already been criticised for allowing too much room for interpretation in terms of levels and outcomes (see discussion in Brown, et al., 2008; Crooks, 2002). The MoE's strategic policy response was five-fold (Brown, et al., 2008; Crooks, 2002; New Zealand Ministry of Education, 1994).

1. Creation of national curriculum statements for all essential learning areas with achievement objectives aligned to eight levels of progression between beginning of primary and end of secondary education;

2. The development and provision of curriculum-aligned teaching materials and programmes to aid educator content and pedagogical content knowledge, with special emphasis on innovations in literacy and numeracy;

3. The development of a toolkit of assessment resources to support educators in school-based assessment consisting of a wide variety of curriculum-aligned assessment resources, including:

    a. an online Assessment Resource Bank (ARB) for English, Mathematics, and Science;

    b. national exemplars of progression in Levels 1 to 4 for all essential learning areas;

c. teacher participation in the administration and marking of curriculum monitoring tasks at Years 4 and 8 (i.e., National Education Monitoring Project—NEMP); and

d. an electronic Assessment Tools for Teaching and Learning (asTTle) software system

4. Funding of extensive professional development for educators with specific attention to assessment for learning (i.e., Assessment for Better Learning - ABEL and Assess to Learn - AtoL)

5. A low-stakes educational assessment policy distinguished by:

   a. No central reporting of school data

   b. No compulsory nationwide testing programme in primary school

   c. Giving choice to educators as to which assessment resources or methods to use

   d. Aligning assessment resources to curriculum objectives and levels

These actions taken together created a climate of increasing trust, confidence, and competence in which national monitoring purposes were met by school-based assessments and improvement purposes were clearly fore grounded in the daily lives of teachers and administrators.

The development of the teaching and assessment resources in NZ coincided with extensive professional development opportunities for educators. This included development specifically aimed at assessment practises (e.g. Assessment for Better Learning (ABEL) which was later replaced by Assess to Learn (AtoL) (Brown, 2008; Crooks, 2002). Educators' knowledge of assessment was also expanded by participation in the development of assessments such as those found in the Assessment Resource Bank (ARB), national exemplars, and asTTle and participation in the NEMP matrix sampling of student performance (Brown, 2008; Crooks, 2002).

Complementary to this strong emphasis on school-based formative assessment, the government obtained public accountability information through the Education

Review Office (ERO) (Brown, et al., 2008). ERO is responsible for triennial evaluation of all schools' performances in NZ through a combination of school self-review and on-site inspection visits to establish quality assurance and legal compliance (Fiske & Ladd, 2001). ERO not only evaluates schools, but also acts as a resource to support schools, educators and school governing body members by providing public accessible support documents.

Comparing New Zealand and South Africa

From this brief comparison of the South African and NZ educational landscapes, three relevant aspects are highlighted (i.e., infrastructure, educator knowledge and assessment philosophy). Infrastructure refers to the quality of school equipment, buildings, and technology, while educator knowledge refers to the content, pedagogical, and content knowledge of teachers as well as their professional role and position in society. Assessment philosophy refers to the purposes of and policies for assessment. Table 5.3 summarizes the similarities and differences between the countries. It is evident that the discrepancies between them are large in terms of infrastructure and educator knowledge; whereas, the assessment philosophies are strongly aligned. Nonetheless, there are huge discrepancies in assessment for learning resourcing and outcomes between the two societies.

Table 5.3: Comparison of the South African and New Zealand educational landscapes

	South Africa	New Zealand
Population *Estimated 2008 mid year population (World Bank, 2010)*	48,687,000	4,268,900
Schools	2,410,501 learners attending 26,099 schools in 2007 (Department of Education, 2007a)	766,379 learners attending 2,593 schools in 2007 **(Data Management and Analysis Division, New Zealand Ministry of Education, 2008)**
Infrastructure	*Variable provision* Well developed in urban, affluent areas Poorly developed and maintained in rural and low socio-economic status areas.	*Advanced provision* Equitably distributed High Technology based
Educator knowledge	*Developing* Many educators had their training under the apartheid system and have received poor training. The professional development system is still immature and experiencing developmental difficulties. There are however a small group of educators who have received good pre-service and in-service training.	*Advanced* Well established pre-service training and professional development system.
Assessment	*Aimed at promoting learning and monitor progress* Centralised and mandated assessment for the national senior certificate in Grade 12 (75% of marks). 25% of the final result is based on school based assessment.	*Aimed at promoting learning and monitor progress* Centralised and government mandated assessments for the final years of schooling from Years 11 -13.

AsTTle in New Zealand

The Assessment Tools for Teaching and Learning (asTTle) software, which is the focus of this paper, is the single most expensive (i.e., more than NZ$17,000 000 ≅ R89,000,000 between 2000 and 2008) addition to the assessment policy and resource-base or toolkit in NZ. AsTTle is an appropriate choice since the NZ government has invested substantially in the development and deployment of a national curriculum- and normative- referenced educational resource making use of advanced computer technology. Indeed, the asTTle software has been hailed as the “best new education investment this government has made” by the NZ Minister of Education, Trevor Mallard (Atlantech, 2003, p. 1). More importantly, the development and use of asTTle has been seen as a solution to

the negative effects of compulsory national testing, while meeting accountability requirements (Hattie & Brown, 2008).

The asTTle system has been described extensively elsewhere (Brown, et al., 2008; Crooks, 2002; Hattie & Brown, 2008; Hattie, Brown, & Keegan, 2003) so only a brief overview is given here. AsTTle provides the autonomous, decentralised schools of NZ an educational technology resource that provides data for school, classroom and learner improvement by assessing student performance in reading, writing, and mathematics in either English or Maori. Since 2002, schools have been provided, upon request, with the asTTle software free of charge and usage is voluntary. The asTTle software allows schools and teachers to create curriculum-aligned customised, standardised, 40-minute tests of mathematics, reading, and writing from large banks of calibrated test questions for English and Maori medium learners. Reporting is against both the objectives and strands of Curriculum Levels 2 to 6 and norms for students in Years 4 to 12.

All asTTle items and tasks were mapped by teachers, content area experts, and curriculum experts according to the NZ curriculum statements for the relevant subjects. Additionally, all items are mapped to a cognitive processing taxonomy (i.e., the Structure of Observed Learning Outcomes—SOLO) in order to categorise student performance on the various tasks according to broad levels of current functioning (Hattie & Brown, 2004). The test-users can select from a suite of graphical reports (including an online catalogue of curriculum-aligned teaching resources) that allow interpretation of the performance of individuals and cohorts relative to norms, standards, and objectives. The various reports were designed and evaluated to meet educational improvement and accountability purposes (Hattie, Brown, Ward, Irving, & Keegan, 2006).

While asTTle depends on computer technology, its development in NZ has been gradual and consistent with the infrastructure available in schools (Hattie & Brown, 2008). Currently, asTTle is in its 7th generation, which permits on-line testing, computer adaptive testing, and longitudinal tracking and goal setting (Hattie, 2009). At the time of this study, asTTle Version 4 was current (Hattie, et al., 2004a).

5.2.6.2 Sampling

The populations accessed for the NZ case study included:

- Principals, teachers and learners using asTTle
- NZ Ministry of Education officials working with asTTle
- Members of the asTTle team
- Professional development providers and school improvement organisations working with asTTle

From these populations, purposeful sampling for the case study was used to ensure that the perspectives of varying stakeholders on the use of asTTle was included. Participants knowledgeable about asTTle were selected in order to provide rich sources of information. Both the outsider perspective and the choice of rich sources of data contribute substantially to the validity of the case study. The following participants[33] were sampled for the study:

Table 5.4: Participants in exemplary case study

Population	Sample
School Users	• Two teachers (Ms Z and Mrs X) and three Year 8 learners from a school that adopted asTTle in 2006. The teachers were identified by their principal as effectively employing asTTle in teaching and learning and the learners have been exposed to asTTle during 2007. • E, principal of a school, who did an evaluation of asTTle and has experience both of using asTTle and the professional development associated with asTTle.
Ministry of Education	• Two assessment division representatives at the NZ Ministry of Education (NZ MoE) – Y and N. It should be noted that the opinions expressed by these officials does not necessarily represent the official Ministry policy.
asTTle Development Team	• Professor J, the creator of the asTTle tool. • Dr H, one of the asTTle developers, who had been a secondary school deputy principal and chief examiner for mathematics before joining the asTTle team.
Professional Developers and Researchers using asTTle	• B, a professional development provider who was involved in the development of asTTle itself and therefore has a profound content and knowledge about the tool. • Two associate directors of Research Centre G (A and D) and Q, a project researcher at the same centre. The centre focuses on research-based educational interventions specifically with Maori and Pacific Island children. Researcher Q uses the asTTle tool in working with schools on writing interventions.

[33] Attempts have been made throughout to protect the identities of the participants. In some cases, the identities of the participants may still be gleaned due to the very specific nature of their knowledge of the asTTle system. These participants have all been briefed on the limitations to confidentiality and have even agreed to have their identities published. All participants received copies of the analysis products and chapter in order to review it prior to publication and to re-evaluate their willingness to have the data published.

The data collection procedure for this sample is discussed in the next section (Section 5.2.6.3).

5.2.6.3 Data Collection

Data for the case study were generated through semi-structured interviews with NZ stakeholders in the asTTle system. The focus of the interviews was on how asTTle was used and the extent of use as well as the factors that facilitate or hinder use of the feedback system. The full interview schedules, transcripts of all the interviews and analysis products are available on the audit trail DVD. The interview data were supplemented with documents collected in NZ.

5.2.6.4 Data Capturing

Interviews were digitally recorded and transcribed and all field notes were captured electronically. Original digital recordings and notes were archived and will be kept at the University of Pretoria for a minimum of 15 years. Recordings were not published on the audit trail DVD, as the publishing would violate confidentiality.

5.2.6.5 Data Analysis

The interviews were analysed using what Saldana (2009) terms 'pragmatic eclecticism', which means that the researcher keeps herself open during the initial data collection and coding to determine the most appropriate methods of coding. A number of First Cycle coding methods (preliminary coding methods) were combined with Second Cycle coding (categorical, conceptual, and/or theoretical organisation). During the First Cycle, coding and recoding, data were analysed according to meaningful units of text, with codes generated through an inductive process and allocated to each unit individually. The First Cycle Coding methods included the following types:

1. Attribute coding: Coding of descriptive and context data captured at the beginning of each interview.
2. Holistic coding: Coding of basic themes or issues, preparatory to more detailed coding, large portions of text are coded with general codes.
3. Descriptive coding: Coding of the topic, smaller units than holistic coding are coded, in this study descriptive sub-codes were included for greater specificity.

4. Initial coding: This coding is also known as open coding, which involves the breaking down of the data into discrete parts and examining them for similarities and difference. This open-ended method allowed for deeper reflection on the data and the inclusion of nuances.
   (Saldana, 2009)

Once the First Cycle coding was completed, codes were clustered in meaningful groups to generate themes. For example, the theme 'professional development' includes codes such as *use to increase data-literacy* and *use to transform data into action*. The Second Cycle approach employed in this study was therefore pattern coding which is both inferential and explanatory, pulling large amounts of codes and data into more parsimonious units, sometimes known as meta-codes (Saldana, 2009).

This type of analysis employed in this thesis is based on inductive logic where the researcher has to immerse herself in the data in order to allow the themes to be constructed. The analysis was aimed at organising, describing and interpreting the data by trying to identify patterns or themes and constructing a framework through which this essence can be communicated meaningfully (Best & Kahn, 2006).

This analysis was framed in the pragmatist paradigm and was underpinned by the conceptual framework for this study (See Chapter 4), therefore the researcher approached the work with some pre-existing ideas. Readings on feedback of monitoring and pre-conceived ideas on the topic influenced the data analysis. The researcher moderated the inherent subjectivity through reflexive notes and memos, as well as the use of peer debriefing and examination of the audit trail by some of the expert consultants (See Chapter 4 for further discussion of the methodological norms for this study). The qualitative data analysis tool *Atlas.ti* was employed to facilitate analysis and provide an easily accessible audit trail in web-page format.

Computer Aided Qualitative Data Analysis

Many tools for Computer Aided Qualitative Data Analysis (CAQDAS) exist. *Atlas.ti* was utilised for the analysis here as it falls in the category of code-based theory building packages (Lewins & Silver, 2009). *Atlas.ti* allows for the analysis

of textual, graphical and audio data (Scientific Software Development, 2004). Willig (2001, p. 151) describes *Atlas.ti* as moving beyond mere coding and retrieval with several additional features including: "…visual displays of the hierarchical relationships between codes and the construction of conceptual diagrams or networks".

> The use of computers can certainly speed up the process of data exploration. The easy retrieval of data files and inspection of analytic memos on screen removes the need for physical cutting and pasting, photocopying of extracts, colour coding and manual sorting. In addition such programs allow the researcher to retrieve files that share certain features based on codes, keywords or descriptive labels in order to identify patterns within the data without having to search through the entire data set. (Willig, 2001, p. 152)

Atlas.ti facilitates the use of direct quotations to enrich the data representation. The use of CAQDAS is specifically indicated when dealing with large amounts of unstructured textual material, which could cause serious data management problems (Henning, Van Rensburg, & Smit, 2004, p. 129).

Separate analytical or hermeneutic[34] units were created for the analysis of the qualitative data in the different design cycles. A unit was generated for the analysis of the case study data from NZ. The separate units will allow for analysing the qualitative data from the various cycles on their own as they relate to different aspects of the design research process.

As *Atlas.ti* can be a complex tool for a novice user, and represents only one of a number of CAQDAS programmes that are available on the market, it may limit how well the reader can interact with the data and analysis using *Atlas.ti*. All data-analysis products and memos relating to the analysis process are therefore presented in the format of web pages for each hermeneutic unit and are included on the audit trail DVD along with this thesis. No knowledge of *Atlas.ti* is required to access the audit trail in html format and the data and analysis products can easily be accessed by any reader.

[34] The units that are used to analyse data in *Atlas.ti* are known as hermeneutic units.

5.2.6.6 Discussion

From the analysis, it appeared that the success of the feedback system incorporated in asTTle in part relied on the structure of the feedback system and its components. The key components contributing to the feedback system are illustrated in Figure 5.8.

I. asTTle Instruments

The first component is not necessarily part of the feedback system, but facilitated the trust in the data and instruments used to collect the data. One of the NZ MoE requirements for the development of asTTle was that teachers be involved in the development and testing of items for the item banks.

> *[Trevor Mallard] made some demands of us and he said 'you must involve teachers', he said, 'I'm not giving you all this money only to find it all going to the university and not going back to teachers'. (PD3, Prof J creator of asTTle[35])*

Teachers who contributed were compensated for their services and acknowledged in the technical documents and instrument manuals of asTTle. The participation in the development of asTTle de-mystified asTTle for many teachers and helped to build trust in the instruments. Great care was also taken in developing the items for asTTle to insure that a lot of the content of the items specifically related to NZ: "*we would be very careful to have material relating to New Zealand in asTTle" (PD1, PD provider B).* This made it more difficult for asTTle to be adapted to other contexts, as these items would have to be replaced with others specific to the new context. However, the NZ specific content engendered more trust and a greater sense of ownership of the instrument for NZ teachers and schools. The instruments for asTTle are highly dependent on ICT and schools design and print the assessments themselves based on the aspects they wish to assess. They set parameters according to the learning area they

[35] Reference to interview document from which quote originates. E.g. PD3, Prof J creator of asTTle, refers to Primary Document 3, Professor J creator of asTTle. The reference can be checked against the documents on the audit trail DVD.

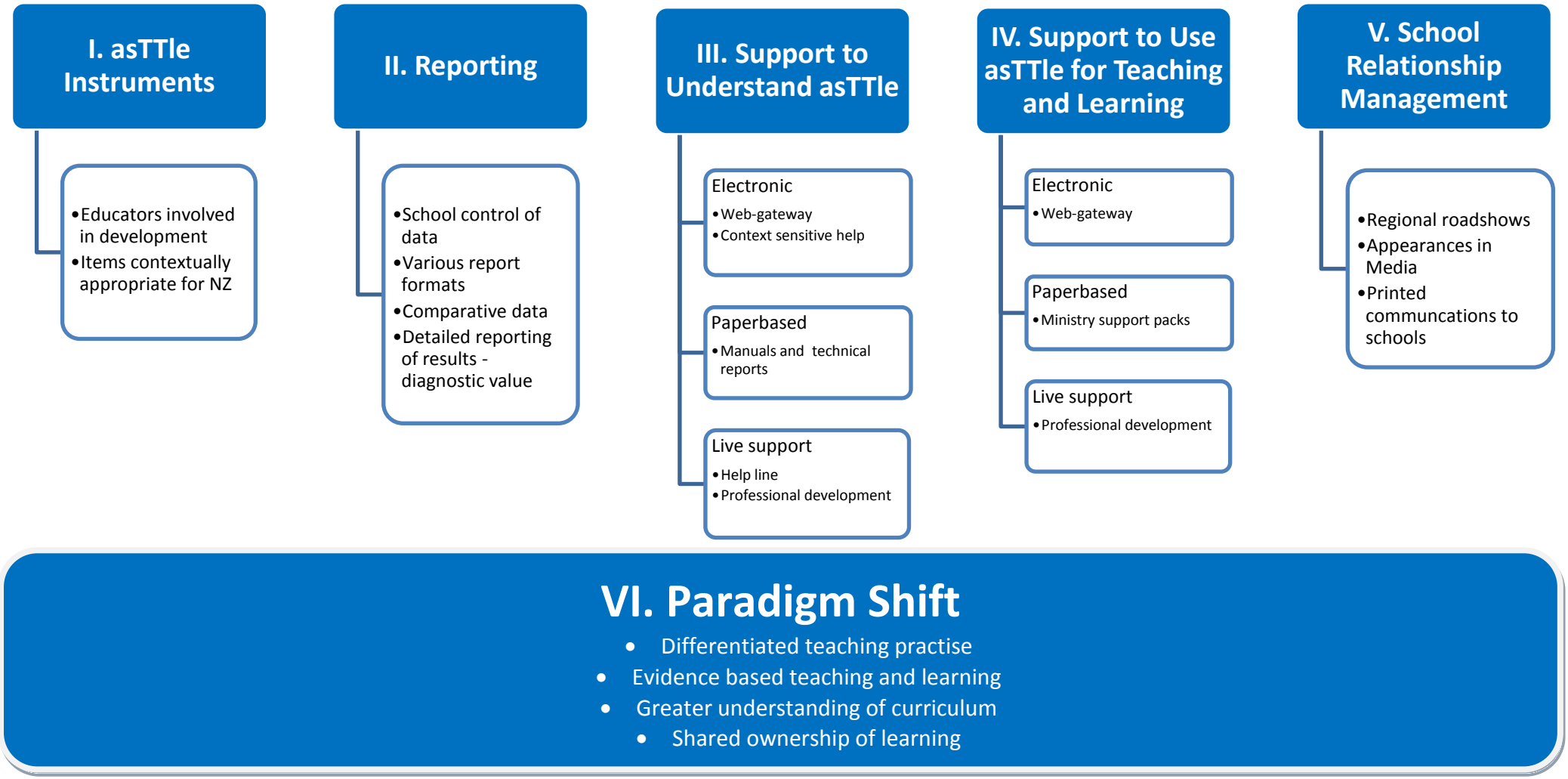


Figure 5.8: Components facilitating the success of the asTTle SPFS in New Zealand

wish to assess, at what level and also the specific topic areas. The assessment is then produced along with a memorandum. All the materials are then printed by the school and each learner receives an assessment.

The teachers then mark the assessment according to the memo and input the data into the system. As the system is so reliant on ICT, the design team was very concerned with ensuring that the technical aspects were not so complex that it detracted from the real purpose of the assessments to support assessment for teaching and learning. "*Absolutely. As I said before … it is all about teaching and learning and if they talk about the technology we have failed" (PD3, Prof J creator or asTTle).*

II. Reporting

The second component relates to the reporting of data. AsTTle can be used by schools autonomously. Schools input the data themselves and can then produce a variety of different reports based on it, which means that the turn-around time from assessment to feedback is very short and solely dependent on the schools. Schools do not have to report data to the ministry, but if they are willing to do so, it is done anonymously. This was one of the pre-requisites of the asTTle developers, that there should be no way in which the data could be used for high-stakes accountability practices by the NZ MoE.

> *I take the view that the government has a need and a right to know, but we don't want league tables, and we don't want teaching to the test, and we don't want the teachers having the test define the curriculum. (PD3, Prof J creator of asTTle)*

The reports were such an important concern for the asTTle developers, that various report formats were part of the proposal to the NZ MoE for the project and the various report formats were piloted prior even to the completion of asTTle, to ensure they were understandable and usable for school improvement.

> *… a lot of focus group time was spent with teachers presenting different versions of reports, asking them what they think it means getting in to tell us ummm…what they needed additionally to improve it, or if it had this if it would make any kind of a difference. (PD5, Dr H asTTle team)*

The report formats present the data in a variety of ways, either on an individual learner, class or school level with comparative data from the rest of the country or rest of the school *"That's the New Zealand mean. And that is where we are up to on the asTTle reading score" (PD7, Year 8 learners).* Comparative data for specific ethnic groups can also be inserted and reports are also presented in a variety of formats ranging from traditional graphs, and tables to those using graphics similar to an odometer to show progress and relative performance. The variety of formats was essential to ensure that teachers and schools could acquire information in a format detailed enough to allow its use for diagnostic and improvement purposes. The different report formats also cater for the preferred style of data presentation of teachers and schools. One of the report formats is also specifically geared to grouping children according to learning needs, to make it easier for teachers to use differentiated learning in their classrooms. These groups are also seen by teachers and learners as not being static, but as based on current learning needs in a specific learning area e.g. Geometry.

> *"[I]n some things the child may be in the purple group and they just pick it up so easy, so you just move them on. The groups are fluid… they will probably change about six times in a term, easily…" (PD10, Ms Z teacher)*

Children are placed in different groups for each learning area and the groups may change and shift as children progress through each learning area. The reports also directly relate back to the curriculum and the areas being assessed to help target action for improvement. *"[T]he reason I prefer asTTle to any of the other tools is that it is linked to the curriculum" (PD1, PD provider B).*

The third and fourth components of the asTTle feedback system relate to reporting, but there are however two distinctly different aims in the support structures. The one focuses on providing support to understand how to use the asTTle system and understand the data. This encompasses the technical skills to use the asTTle system and to understand the reports and graphs in asTTle. The other focuses on how to move from understanding the data and

system to using it for curriculum and classroom planning, teaching and learning activities.

III. Support to Understand asTTle

The support to understand asTTle is provided in a variety of forms. First, there is electronic support in the form of the www.tki.nz.org gateway that provides links to technical reports and manuals for asTTle as well as information pages. The programme itself is also equipped with a context-specific help function that facilitates the technical use and understanding of the programme. For the less technically inclined, paper-based instrument, manuals and technical reports are also available. Live support is also in place in the form of a helpline to the asTTle team at the University of Auckland. If the call centre personnel cannot assist the calls are immediately escalated up to members of the asTTle development team themselves. Schools can also contract professional development (PD) providers to provide training and support for the use of asTTle.

> *Well, at the moment our strategy is based on assessment for learning, so our professional development is helping teachers use assessment for learning better in their classrooms and schools. Part of that is purposeful use of tools. The tool part is a small part of the work we are doing. (PD 8, N NZ MoE).*

IV. Support to Use asTTle for Teaching and Learning

The fourth aspect, support to use asTTle for teaching and learning overlaps somewhat with support to understand asTTle. The focus here, however, is not technical, but school improvement, moving from understanding what the data is relaying to interpreting it and taking action accordingly in a specific school or classroom. This aspect builds on the support to understand asTTle. The TKI gateway provides electronic support to link the data to action through a 'What Next Button'. This function helps to relate the data to what the next area of learning should be and even has examples of activities that may help to achieve the learning. *"[W]hen I go on to asTTle 'What Next?' it is absolutely what I want to meet the learning needs. It is not just lots of cute activities" (PD 9, Mrs X specialist teacher)*

The NZ MoE has also produced information packs about asTTle:

> *So we produce the technical reports, and that is very important, but we are also producing information that is far more accessible for policymakers and for teachers. So for example in this one, this was a tool that we had that looked at reading, maths and writing [referring to asTTle data]. (PD8, Y NZ MoE)*

These not only discuss the project and introduce it, but also provide guidelines on how to interpret data and relate it back to action in a specific school using examples and guidelines. PD providers can also be contracted for training on transforming the asTTle data into action in schools. PD for asTTle not only addressed teacher technical requirements, but also emphasised the underlying assessment *for* learning philosophy of asTTle (Hattie, et al., 2006).

V. School Relationship Management

A final part of the feedback system is a relationship management function. Part of the trust that asTTle users have in the system is related to its creator Professor John Hattie being a well know educational researcher in NZ.

> *Having Prof J as the programme director has meant a huge amount of trust involved, because he is such a respected figure in the field of education…and having somebody as respected as that has made it a lot easier" (PD1, PD provider B).*

The project team are encouraged to engage in educational activities throughout NZ and upon occasion to appear in the media to comment on educational matters. This contributes to the visibility of the project and the trust that users place on asTTle. The MoE also regularly sends out printed information about asTTle and developments to keep users informed of developments. Annually the asTTle team also undertake a number of regional meetings to meet with schools. Such meetings provide an opportunity for direct communication with schools about their context, problems and experiences with asTTle. The meetings are invaluable in putting a human face to asTTle.

This interaction of various components of the asTTle feedback system helped overcome initial resistance to asTTle and brought about a number of paradigm

shifts. Originally, there was a theme of ***Resistance to the introduction of asTTle***. This particularly related to:

1. *Fear of high-stakes accountability.*

As asTTle was funded by the NZ MoE many teachers were afraid that it would be used to monitor teachers through high-stakes accountability practises.

> *[A] lot of people thought it would be used as a tool of torture and that it was going to be used to compare school A against school B. It was going to be part of a national monitoring system where schools would have to give the information to the ministry… (PD2, Principal E)*

Interestingly, the developers of asTTle at the University of Auckland were vociferous opponents of such use of asTTle (Hattie, 2005).

2. *Wide range of assessment tools available in NZ.*

AsTTle is only one of a range of government-sponsored or approved assessment resources available to NZ schools (Brown, et al., 2008; Crooks, 2002). Schools are free to employ any of the tools that they feel are appropriate, though unfortunately many have not lived up to expectations (e.g., they are expensive, time-consuming, or technologically cumbersome) and teachers were hesitant to accept the promises about asTTle based on previous experiences with other promising new assessments.

3. *Vulnerability through transparency.*

The asTTle reports are geared towards transparency in assessment. The results are diagnostic and individualised for each learner. *"[AsTTle] gives us the information we need, is informing the practice and is very easy to share with the students at this age." (PD9, Mrs X specialist teacher).* This means a high level of openness about learner performance, which some educators and school-leaders may find threatening. Principals and HoDs also have greater insight into the comparative performance of various classes in the school.

4. *Shock of standardised results compared to national standards.*

Given the validity procedures asTTle employed (i.e., independent curriculum mapping, national teacher item review panels, and national teacher standard setting panels) it is highly likely that asTTle curriculum level reports reflect appropriately the objectives of the national curriculum statements. Nevertheless, gaining credibility in the minds of teachers was more complex. Many educators were shocked at the first results from asTTle (Hattie et al., 2003), believing that asTTle had given lower curriculum level scores than students merited. For example, teachers may have perceived their learners to be at Level 4, while the asTTle results may have revealed their learners were only functioning on Level 3. *"[A] lot of the teachers think these kids are doing well, and then the test is saying they're not doing well." (PD5, Researcher D).* Thus, there was a tendency on the part of educators to question the validity of the asTTle tool, rather than examining the standards they were employing in the classroom*. "[R]ather than having a good look at their understanding of Level 4 and whether the students are fulfilling, grasping the concept of that level, they were criticising the tool" (PD2, Principal E).*

5. *Fear of change and the novel.*

Many educators are entrenched in their assessment practices and were hesitant to try a new approach. As one teacher put it: "*it would be a problem for teachers who were entrenched in 'this is the way I have done it for the last 104 years, it is working, why change it?'" (PD9, Mrs X specialist teacher).*

6. *Technological complexity.*

Although NZ has a very good ITC infrastructure (e.g., all schools in NZ at the time of asTTle V4 deployment in 2005 had access to broadband internet connectivity of at least 54Mbps; all teachers were entitled to participate in a national government-funded "laptops for teachers" scheme) and most educators are computer literate, some teachers found asTTle challenging. Some NZ educators were also not comfortable working with computers, which dissuaded them from employing the system; *"So teachers do get terrified from it, some of them do get terrified because it is computers" (PD3, Prof J creator*

of asTTle). Indeed more positive attitudes towards ICT predicted greater usage of asTTle, which indirectly contributed to more accurate interpretation of asTTle reports (Hattie et al., 2006).

7. *Cost.*

Printing of asTTle assessments and reports increased stationery expenditure in schools; instead of buying off-the-shelf test products schools were printing their own customised tests. *"One principal said to me the other day, it is a dollar (sic) for each paper, which is quite a lot it is $30 for a class. So, that is a barrier for some people." (PD1, PD Provider B).* Furthermore, construction of assessments, capturing of marks and printing of reports could also be time consuming, particularly if a person were unfamiliar with the programme. *"[I]t takes a lot of paper, it takes a lot of time for the teachers to do the marking and the data re-entry" (PD5, Dr H asTTle team).* Thus, some schools found these factors outweighed the benefits of asTTle. Nonetheless, it could be argued that this was a matter more of changing teachers' work rather than adding to it; the computer saved considerable time to create and analyse tests (e.g., creation of photocopy ready tests took 10 minutes in asTTle, unlike the four or more hours needed to prepare tests even when items were cut-and-pasted from previous tests).

A number of themes related to ***how asTTle overcame the resistance to implementation:***

1. *Aligned development of infrastructure, educator knowledge, and assessment philosophy.*

At the time of deployment, schools had in place the necessary infrastructure of ICT (including copying facilities) so that usage of asTTle was feasible. However, asTTle was presented not as a technology or testing resource, but rather an educational resource (Hattie & Brown, 2008*). "[W]e will never talk about testing, we will talk about teaching and learning- how is it going to make a difference to teaching and learning" (PD3, Prof J creator of asTTle).* Furthermore, asTTle was introduced in a policy environment that strives towards assessment *for* learning (New Zealand Ministry of Education, 1994)

and enhanced teacher quality. The Ministry provides teachers various forms of PD (e.g., assessment, literacy, numeracy) to reinforce this approach and their pedagogical content knowledge. Therefore the PD, electronic and paper-based support were all aligned to the underlying philosophy of assessment for learning.

2. *Support for use.*

This includes the electronic, paper-based and live support already discussed above.

3. *Tools to inform change.*

AsTTle differentiated itself from other assessment tools in that it provided highly diagnostic information compared to both curriculum expectations and national norms. Teachers can access reports for individual learners, grouping reports, reports comparing class progress against national norms, and so on (Hattie & Brown, 2008; Hattie, et al., 2003). The data are also specific in terms of the differential skills of learners in different learning areas. "*I had a child last year who on reading tests would actually do really well, but on asTTle it actually showed up that she had a gap in evaluation, she was actually quite low" (PD9, Mrs X specialist teacher*). This information enabled the teacher to identify what even better children needed to learn. Results are operationalised in terms of gaps, strengths, weaknesses and things still to be learnt. AsTTle then also provides links to what actions can be taken in the class to move learners from their current levels to the next level.

4. *Trusting intentions.*

From the beginning, the terms of development for asTTle made it clear it was for school-based use, not government or central agency use. Data from asTTle are not sent to the MoE (either by regulation or by technological means), although schools may choose to use their data to demonstrate their effectiveness to the Ministry in their triennial reviews. This explicit policy of "no control, no compulsion, no central reporting" (Hattie & Brown, 2008) helped to break any distrust schools may have had towards the introduction of a

government-funded national assessment tool. The voluntary status of the tool also mattered to teachers: *"this is not a country where teachers take kindly to being forced to do things" (PD1, PD provider B).* That the instrument was designed by a team outside the government and led by a well-known academic figure (Prof J) further engendered trust in the tool and its intentions. As part of the Ministry contract, teachers were involved in the development process of asTTle and in this way teachers could see the inside workings of the design of asTTle and contribute to it. These teachers later became resources in spreading information on asTTle, either through formal avenues such as PD, or by word-of-mouth. *"Well, people say to me, well how did they write these questions? And I say, well I know exactly the answer to that question, because I helped put it in." (PD1, PD Provider B).* Indeed, the point of assessment for learning is to encourage schools and teachers to discover weaknesses or poor results in their students' learning without fear of shame, blame, or punishment.

5. *Success in bringing about learner improvement.*

AsTTle has in many instances formed the basis for drastic improvement in learners. Once educators could identify the learning needs, these could be addressed and shifts in the learning could be accurately measured against the curriculum using asTTle. *"[T]hat is just in a term, the difference you can see from just a term, so teach it for a whole year, like it is just exciting to see …" (PD 10, Ms Z teacher).*

VI. Paradigm Shift

The most fundamental changes brought about by the asTTle feedback system were shifts in paradigm in schools, PD providers as well as academic and MoE staff. These shifts were stimulated by the high degree of congruence between the underlying philosophy of the feedback system and the educational policy context. These paradigm shifts included:

1. *Using assessment for evidence based practice.*

While NZ has no compulsory testing in primary school, there was a great deal of testing, much of which was seen as producing results for records. In contrast, the asTTle system employs a formative, validity-focused process that begins with teachers determining the focus of any asTTle assessment, followed by diagnostic interpretation of data directly linked to teacher action to help learners to progress to the next level. The tool provides a basis for evidence-based classroom practice that is sensitive to the needs of the learners.

> *[F]or years, we were sort of feeling around in the dark, this might or might not work, and now we have some clear evidence to show where our problems are. We can target our teaching skills" (PD1, PD provider B).*

In some cases, the introduction of asTTle led to greater transparency around assessment with learners and parents being briefed on results and their meaning for further teaching and learning. *"I use [the reports] with parents and I say here is the data, this is where the child is at and what we are working towards" (PD10, Ms Z teacher). "[AsTTle] gives us the information we need, is informing the practice, and is very easy to share with the students at this age" (PD9, Mrs X specialist teacher).*

2. *Shared ownership of learning.*

By shifting the assessment practices from a purely summative or teacher-centric process to a more assessment *for* learning approach, learners were provided the opportunity to take responsibility for their own learning. *"It is a little like doing a dance with them, we are in it together… I think the engagement, self-monitoring, the talking with them [learners] makes a difference" (PD9, Mrs X specialist teacher). "[I]f you don't work for it, you are not going to get it, so you are in control of your learning" (PD7, Year 8 learners).* Openness and sharing of reports from assessments allowed learners to gain insight into their current status and what their needs were and thus motivating them to focus their efforts.

3. *Facilitation of differential teaching practises.*

AsTTle provides various reports, including grouping reports, where learners with similar needs are grouped together. Teaching material and activities can thus be differentiated for the group's particular learning needs. Grouping in such a manner means that learners do not associate groups with specific levels, but differing learning needs. No negative connotations of a 'clever' or 'stupid' group are thus created. "*[I]t's never like it is in junior school… like when you get the best readers in the top group. It's like what you got wrong, it is what you don't understand." (PD7, Year 8 learners)*

4. *Increased understanding of curriculum levels and description.*

The autonomy of NZ schools and the breadth of curriculum outcomes have jointly facilitated a divergence of understanding between teachers as to expected standards. Teachers tend to be exposed to learners from a certain region and socio-economic background and naturally revert to a form of norm referencing and personal interpretation of the curriculum standards and outcomes. AsTTle, through its writing of items to curriculum objectives and curriculum-levels based standard setting processes, permits the development of a national understanding of progression. *" …[A]ll our learning intentions are given to us in asTTle so we can see what a child at that level, at level 2 should be learning, what their learning intentions should be." (PD10, Ms Z teacher).* This effect was more obvious in the context of written language.

5.2.7 Design Principles from the Exemplary Case Study

The analysis of the exemplary case study proved to be a successful tool to identify design principles for the optimising of the SAMP feedback system in South Africa. The system must include: a trusted assessment system, clear reporting, support to understand the data, support to use the data and school relationship management. The design guidelines for these components are now discussed.

I. Instruments

a. Teacher involvement in development of the monitoring and feedback system is essential to engender trust and ensure contextual appropriateness.

b. The design and building of trust in the assessment is a key component in ensuring the success of the feedback system.

II. Reporting

a. Data must be provided in a variety of formats graphically, textually, tabulated, to accommodate the needs and preferences of different users.

b. Data should already be presented in a way that is clear and easy to understand, the presentation should not demand to high a level of data-literacy from users.

c. Turnaround time from assessment to reporting must be as short as possible.

d. Data must be detailed and have diagnostic value.

e. Comparative data support interpretation and action.

f. Data must be presented in such a way that it allows for additional analysis.

III. Support to understand the feedback

a. Some support must be aimed at facilitating the understanding of the feedback. Making sense of the data as represented.

b. Support should be provided in a variety of formats to suit user needs and preference, possibly including live support, professional development, ICT resources and printed media.

c. Some support should be available around the clock through either printed media or ICT resources.

IV. Support to use the feedback

a. Some support must be aimed at facilitating the use of the feedback must be provided with the data.

b. Support should be provided in a variety of formats to suit user needs and preference including live support, professional development, also ICT resources and printed media.

c. Some support must be available around the clock through either printed media or ICT resources.

d. There must be congruence between the infrastructure and feedback and support delivery modes to ensure sustainability and accessibility.

V. School relationship management

a. Relationship management with users and stakeholders is an essential and ongoing process.

b. Open face-to-face communication with users engenders trust.

c. Some support should be available around the clock, even if only in printed form.

d. The feedback team must be responsive to user input.

e. Feedback facilitator reputation and persona affect the trust in the system.

VI. Supporting paradigm shifts

a. The monitoring and feedback system should exemplify and operationalise assessment for learning to facilitate action in schools.

b. The feedback system should encourage triangulation of results to support evidence-based practise.

c. The feedback system must not be so technical or demanding of data-literacy that the main focus is on trying to understand the system rather than using it.

These design guidelines along with the guidelines from the literature review (Section 5.2.5) and school inputs provided the guidelines for the components and underlying philosophy for the optimising of the SAMP feedback system. Guidelines about reporting and instruments were incorporated immediately as the initial prototyping cycles focused on creating conditions for the use of the data. Guidelines about support and aspects of school relationship management were only fully implemented and explored later in the optimisation of the feedback system as the focus shifted to facilitating the use of the feedback provided.

5.3 Conclusion

In this chapter the pre-existing feedback system was introduced. The context and needs analysis as well as the first cycle of the exemplary case study of asTTle the NZ SPFSs were explored. The combined design guidelines from the Preliminary Phase are summarised in Table 5.5.

Table 5.5: Combined design guidelines from the Preliminary Phase

Component	Case Study	Literature review
Instruments	Use of school and educator expertise in development, to ensure contextualisation and engender trust.	Use of school and educator expertise in development, to ensure contextualisation and engender trust.
Reporting	• Use multiple forms of data presentation to accommodate the needs and preferences of users. • Data presented in clear and easy manner, not require high level of data-literacy • Short turnaround time from assessment to reporting. • Detailed diagnostic data. • Comparative elements. • Allow for further independent analysis	• Provide comparative data for evidence-based decision-making. • Short turn around time between assessment and reporting. • Use of ICT to decrease turn around time and increase autonomy.
Support to understand the feedback	• Multiple forms to support understanding to suit user needs and preference may include live support, professional development and also ICT resources and printed media. • Some support must be available 24 hours, e.g. ICT resources or printed media.	
Support to use the feedback	• Multiple forms to support use of the feedback to suit user needs and preference • Some support must be available 24 hours. • Congruence between the infrastructure, feedback and support delivery modes to ensure sustainability and accessibility.	• Resources to support school improvement based on the feedback must be provided.
School relationship management	• Essential and continuous process. • Open face-to-face communication with users engenders trust. • Some support should be available around the clock, even if only in printed form. • Must be responsive to user input. • Feedback facilitator reputation and persona affect the trust in the system.	• The quality of interaction between facilitator and users impact on sense of trust, ownership and credibility. • Communication must be honest, open, clear and respectful.
Supporting paradigm shifts	• System should operationalise assessment for learning to facilitate action in schools. • Encourage triangulation of data. • The feedback system must not be so technical or data-literacy demanding	• Not be viewed as part of high-stakes accountability

The data from this chapter were used to provide the overall design guidelines for the Prototyping Phase which is discussed in the next chapters.

CHAPTER SIX
Prototyping Phase: Establishing conditions for use (Cycle 1-2)

> Over the past decade there has been exponential attention paid to accountability, assessment for and of learning, and the glut of information that accompanies such phenomenon (sic), aided and abetted by the fantastic increase in technology. Having reams of data, however, does not by itself make us smarter – only overloaded and confused. (Fullan in Earl & Katz, 2006, p. vii)

Plomp (2009) describes this phase that builds on the problem identification, needs and context analysis as the Prototyping Phase of the design research process. This chapter takes a closer look at the research design and results for the first two of the three design cycles employed during the Prototyping Phase. The design focus of the Prototyping Phase initially was to establish the conditions for use of the feedback system, then progressed to facilitating the transformation of these conditions into use. The first two cycles that focused on establishing conditions for use of the feedback system are discussed in this chapter. The next chapter (Chapter 7) focuses on the third cycle in the Prototyping Phase that examines how to facilitate the use of the feedback in schools.

The two complete research cycles of this chapter consisted of the development of two successive prototypes of the feedback system both of which were formatively evaluated to inform the development of the next prototype. The focus of this chapter is illustrated graphically in Figure 6.1.The pre-existing system and feedback prototypes are illustrated in blue and the evaluation activities are illustrated in green. Every full design cycle consists of the prototype adaptation followed by implementation and corresponding formative evaluation of that prototype.

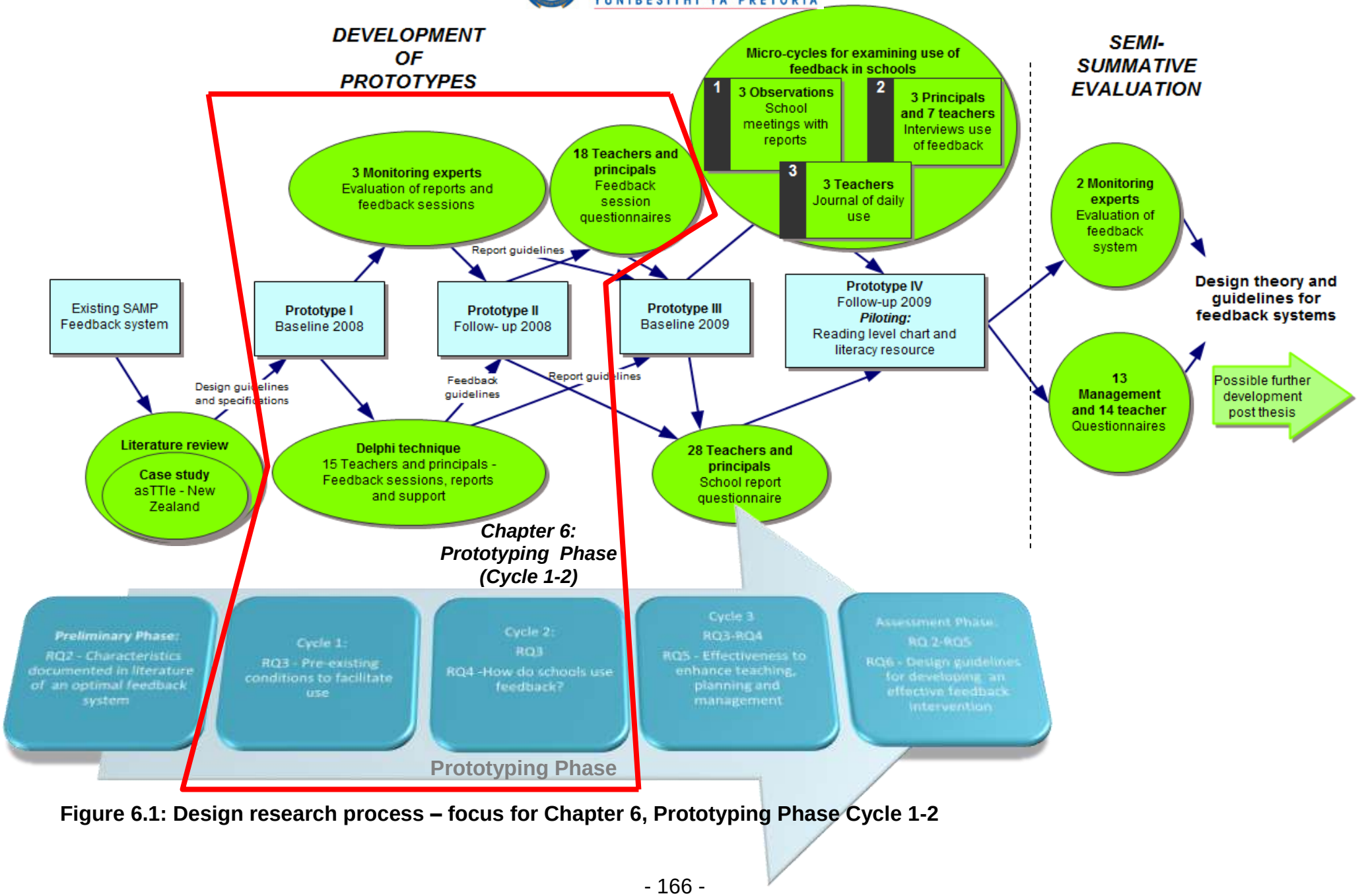


Figure 6.1: Design research process – focus for Chapter 6, Prototyping Phase Cycle 1-2

In the following section each of the three cycles of the Prototyping Phase are discussed separately, beginning with a description of the prototype, followed, by a discussion of the research design used to evaluate the prototype. The evaluation activities are guided by the research questions being addressed and the evaluative focus for the specific cycle. While there is a specific focus for each cycle there is always some deliberate overlap and a cycle may address some aspects of other research questions or evaluative foci. The research procedures for each cycle examine the sampling, data collection and instruments, analysis and discussion.

6.1 Cycle 1 (Prototype I - Baseline 2008)

Feedback Prototype I incorporated the learning from the Preliminary Phase to establish the conditions for use of the feedback. The design and evaluation focus for this cycle was on the reports and feedback session elements of the feedback system.

The formative evaluation for this cycle employed the judgements of both monitoring experts and school users. The monitoring experts were asked to evaluate the reports and feedback sessions. Teachers, HoDs and principals were asked to comment on reports, feedback sessions and support materials. In this case, the monitoring experts and the school users acted as revisors (See Section 4.2.2 for a full discussion of the various roles of participants in design research evaluations), as they were asked to make suggestions for improvement and change. Open, general evaluation questions were employed to allow evaluators to make suggestions about the global design elements of the feedback system, focussing on the feedback session and reports as well as commenting on the monitoring system, logistical factors and basic support and communication with the research team.

Cycle 1 addressed the third ***research question***:

3. *What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?*

The cycle served to provide ***design guidelines*** relating to the design specifications to establish use of the feedback system:

2. *Establishing conditions for use:* This development stage was aimed at improving the components of the feedback system. For this cycle the focus was examining in detail the design of the reports, feedback sessions, logistical arrangements and basic support in detail.

The evaluation in this cycle focused specifically on the ***evaluative foci*** of relevance, consistency and expected practicality (Plomp, 2009), with specific reference to the feedback session and reports:

1. *Relevance (content validity):* The system and its design should be based on state-of-the-art (scientific) knowledge. The feedback system must be clearly connected, in form and purpose, to the learning performance monitoring system for which it provides feedback.

2. *Consistency (construct validity):* The system must be 'logically' well designed. The various parts are well defined and the connections between the parts explicitly postulated. Although some elements may emerge more clearly throughout the process, there are no internal contradictions.

3. *Expected practicality:* The system is expected to be usable in the settings for which it has been designed and developed. Here in particular the focus is on whether the support, feedback sessions and reports are perceived as understandable and helpful for the schools in informing action.

In the following section Prototype I of the feedback system, developed and implemented in Cycle 1, is introduced briefly.

6.1.1 Prototype I – Baseline 2008

After 2006, the mode of assessment for SAMP changed to paper-based in Afrikaans, English and Sepedi, which proved more feasible for adaptation and sustainability (see Chapter 2). This shift meant that the sample size of schools

participating in the SAMP monitoring system was increased to twenty-two (English, Afrikaans and Sepedi) which translated into 1,535 learners being assessed in the baseline for 2008. All the schools in the sample also participated in the feedback system. Data from the assessments were captured by the fieldworkers, using optical mark forms, assessing learners on a one-to-one basis. Data were captured electronically, cleaned and analysed at the CEA, without being sent to the CEM, thus decreasing transposition errors and turnaround time for data cleaning, analysis and reporting.

All the schools in the sample received paper-based reports that included an instrument manual section. The teachers, principals and HoDs involved in the project were invited to attend the feedback session at the University of Pretoria. Schools were also invited to contact the CEA if they required further support interpreting and using the feedback. The components of Prototype I are shown in Table 6.1.

Table 6.1: Prototype I – Baseline 2008 components

Component	Description
Paper-based baseline report	Produced for each school individually and handed out at the feedback session
Baseline feedback session	Principals and teachers from all participating schools invited to the University of Pretoria
Telephonic, written and face-to-face communication	On an ad hoc basis as required

The inclusion of a component of more open communication with schools was based on the guidelines from the previous cycles. Schools were invited to approach the research team directly with any queries or if they required support in any improvement actions. These communications took place on an ad hoc basis as requested by schools. The changes to the paper-based reports and feedback sessions are discussed below.

6.1.1.1 Reports

The overall structure of reports was revised, with the adapted structure of the reports as follows:

1. Introduction to the SAMP project
2. Introduction the assessment subtests and scales (revised and expanded)
3. Results per school (expanded and revised)
4. Individual results (notes on interpretation and histograms added)
    a. Notes on interpretation
    b. Learner results table
    c. Learners at risk
    d. Exceptional learners
5. Conclusion and recommendations (expanded, recommendations added)

The introduction to the project remained similar to that used in the pre-existing system, as did the description of the subtest and scales. An additional section was however added with a figure to represent how the subtests and scales relate to each other (see Figure 6.2).

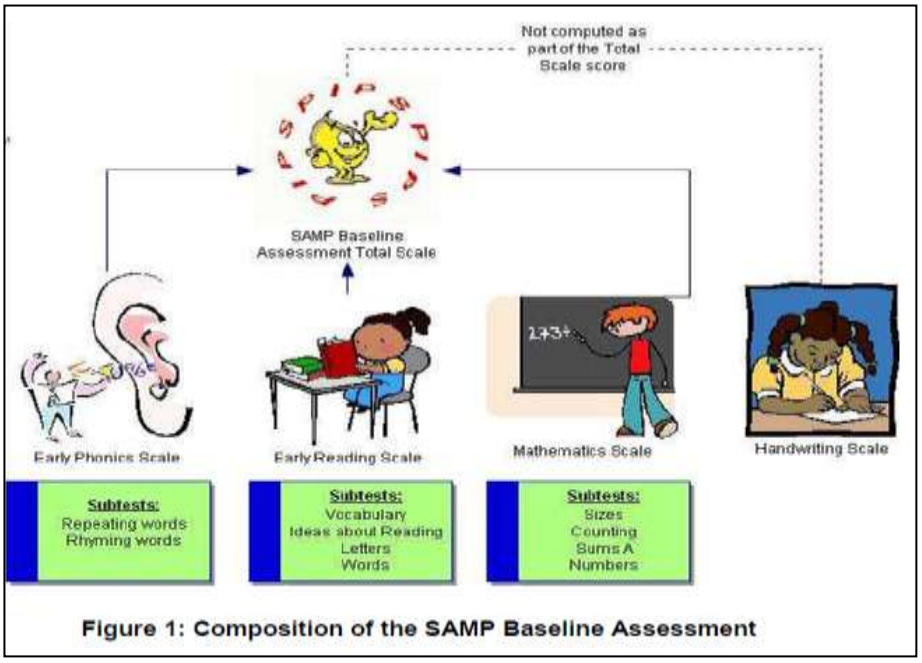


Figure 6.2: Composition of the Baseline Assessment – Prototype I

The graphics from the computer-based assessments reproduced in the reports were replaced by the graphics from the paper-based assessment (see Figure 6.3). All the paper-based graphics had been adapted to the South African context

as discussed in Chapter 2. In the pre-existing system the graphics represented in the report were usually from the easiest items of the assessment. This raised concern that teaching to the test may take place in the classrooms, based on the descriptions and example items. Therefore, Prototype I rather incorporated graphics from equivalent items, or the most advanced items, to maintain the integrity of the assessment.

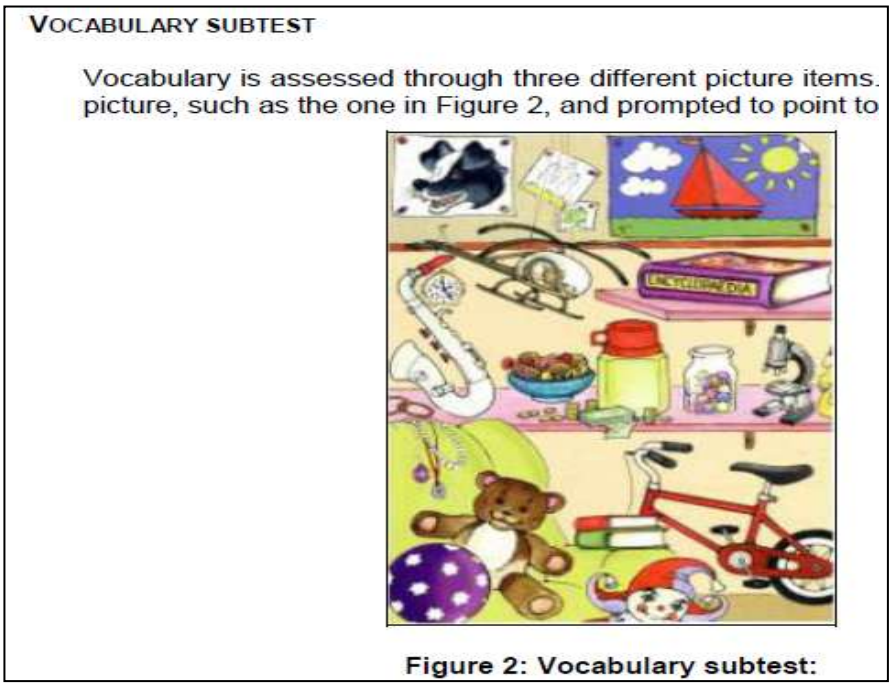
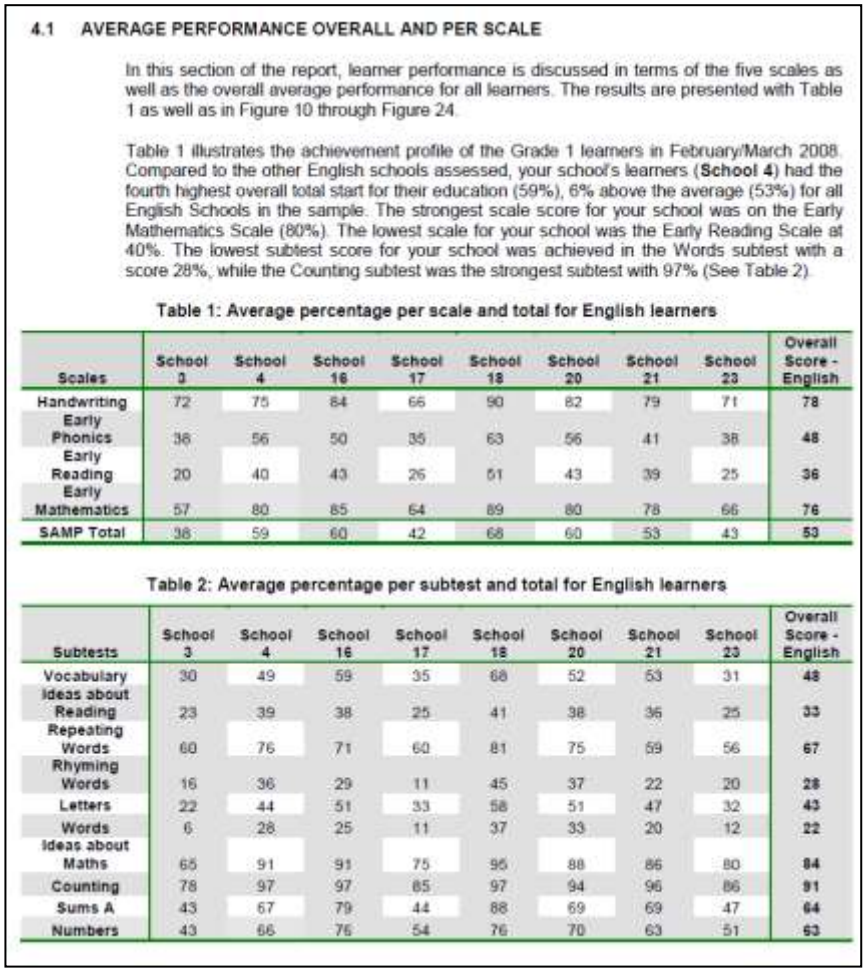


Figure 6.3: Description of the Vocabulary subtest – Prototype I

Section 3 dealt with the overall school results. Unlike the pre-existing reports, all school-aggregated data were represented as percentages and not as aggregated categories to provide more detailed data. As the sample size was increased and the paper-based assessment did not have the technical problems experienced with the computer system, comparative scores were generated for each scale and subtest, whereas previously, they had been limited to the Early Mathematics Scale. Schools were allocated numbers to maintain the comparative component without impinging on trust, anonymity or confidentiality. Results were only compared across schools assessed in the same language, as although the assessments were similar, equivalence had not been established. The section on overall results for the school started with a tabulated summary of results and a discussion to provide schools with a quick overall guide of performance (see Figure 6.4).



4.1 AVERAGE PERFORMANCE OVERALL AND PER SCALE

In this section of the report, learner performance is discussed in terms of the five scales as well as the overall average performance for all learners. The results are presented with Table 1 as well as in Figure 10 through Figure 24.

Table 1 illustrates the achievement profile of the Grade 1 learners in February/March 2008. Compared to the other English schools assessed, your school's learners (**School 4**) had the fourth highest overall total start for their education (59%), 6% above the average (53%) for all English Schools in the sample. The strongest scale score for your school was on the Early Mathematics Scale (80%). The lowest scale for your school was the Early Reading Scale at 40%. The lowest subtest score for your school was achieved in the Words subtest with a score 28%, while the Counting subtest was the strongest subtest with 97% (See Table 2).

Table 1: Average percentage per scale and total for English learners

Scales	School 3	School 4	School 16	School 17	School 18	School 20	School 21	School 23	Overall Score - English
Handwriting	72	75	84	66	90	82	79	71	78
Early Phonics	38	56	50	35	63	56	41	38	48
Early Reading	20	40	43	26	51	43	39	25	36
Early Mathematics	57	80	85	64	89	80	78	66	76
SAMP Total	38	59	60	42	68	60	53	43	53

Table 2: Average percentage per subtest and total for English learners

Subtests	School 3	School 4	School 16	School 17	School 18	School 20	School 21	School 23	Overall Score - English
Vocabulary	30	49	59	35	68	52	53	31	48
Ideas about Reading	23	39	38	25	41	38	36	25	33
Repeating Words	60	76	71	60	81	75	69	56	67
Rhyming Words	16	36	29	11	45	37	22	20	28
Letters	22	44	51	33	58	51	47	32	43
Words	6	28	25	11	37	33	20	12	22
Ideas about Maths	65	91	91	75	96	88	86	80	84
Counting	78	97	97	85	97	94	96	86	91
Sums A	43	67	79	44	88	69	69	47	64
Numbers	43	66	76	54	76	70	63	51	63

Figure 6.4: Comparative tables per language group for scale and subtest scores – Prototype I

The tabulated results were followed by bar graphs of results comparing the school's performance with that of other schools in the same language group. The comparative performances of individual schools on each scale (Overall, Early Phonics, Early Reading, Early Mathematics and Handwriting) were represented graphically, together with a discussion. After the representation of each scale, the graphs for the subtests (that constitute the particular scale) were also produced (see Figure 6.5). For example, the Phonics Scale comparative graph would be followed by comparative graphs of the Repeating Words and Rhyming Words subtests that constitute the Phonics Scale.

4.1.3 PHONICS SCALE

The Phonics Scale consists of the Rhyming Words and Repeating Words subtests represented in Figures 13 and 14. The learners from your school achieved the joint second highest Phonics scale score (along with school 20) of all the English schools who participated in the Baseline assessment with 56%. This is 8% above the average of all the English schools in the sample. Your school performed better on the Repeating Words than the Rhyming Words subtest which comprises this scale (76% and 36% respectively).

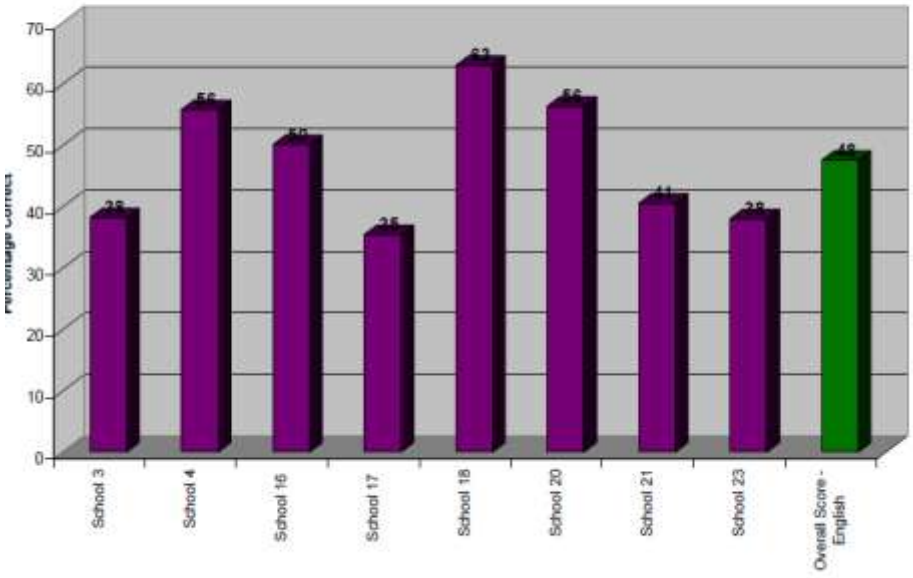


Figure 12: Performance on the Early Phonics Scale per school and overall – English medium of instruction

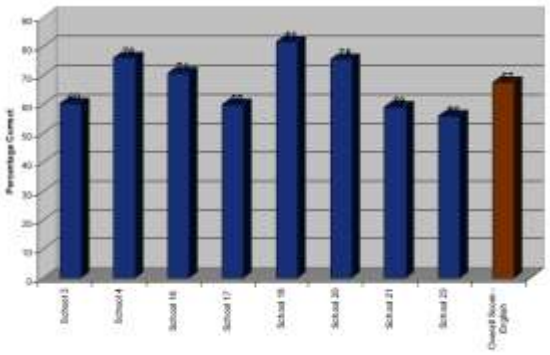


Figure 13: Performance on the Repeating Words Subtest per school and overall – English medium of instruction

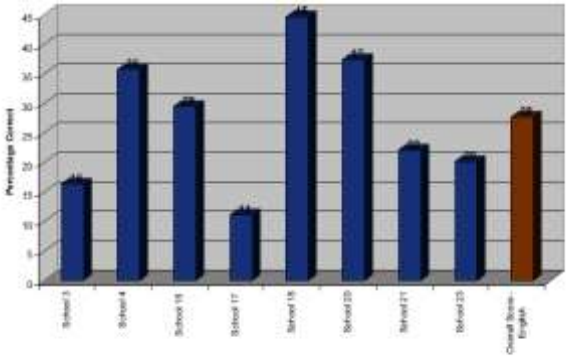


Figure 14: Performance on the Rhyming Words Subtest per school and overall –English medium of instruction

Figure 6.5: Early Phonics scale and constituent subtest results across schools – Prototype I

The overall school result section concluded with a graph summarising the school's performance across the scales (see Figure 6.6), then across the subtest. The tables containing the easiest and most difficult subtest items for each school was also removed from Prototype I as schools indicated that it did not contribute any value and that the descriptions of the subtests and scales were more meaningful.

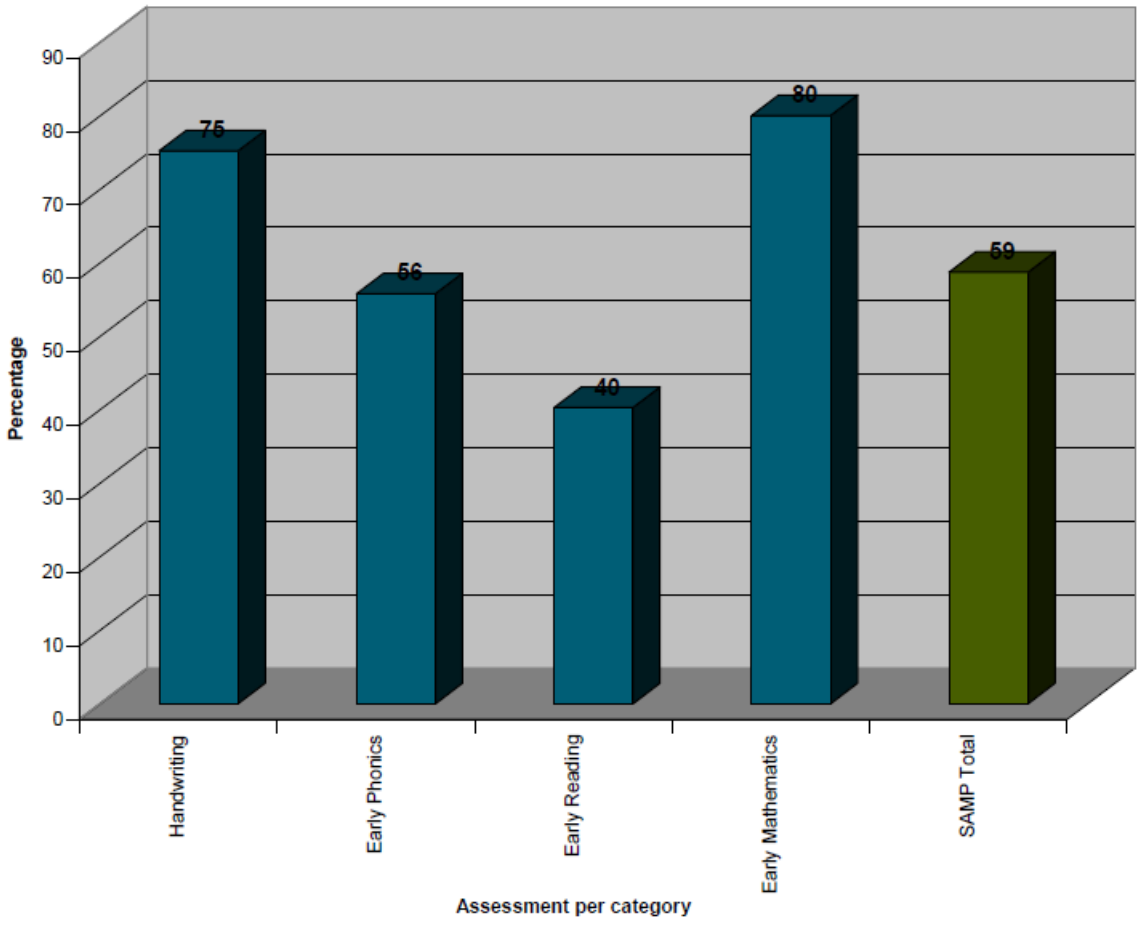


Figure 25: Performance per scale for School X Primary School per scale and overall

Figure 6.6: School's performance across the scale – Prototype I

The Individual learner's results in Section 4 maintained the representation of the data as categories from 1-4. A section with notes on interpretation[36] was added to warn users against over-interpretation of results for a learner from a single assessment. The tabular representation of learner results presented in categories for each of the scales for the baseline was also kept (see Figure 6.7).

[36] It is important to remain cognisant of the fact that all testing situations are dynamic. Children's performances are influenced by multiple factors on any specific test day. A learner's performance may be influenced by such factors as his or her health, the time of day at which the test is administered, whether or not the child is hungry, emotional difficulties or even the rapport with the specific fieldworker.

Table 3: Individual learner results for SAMP Baseline Assessment School X Primary School 2008 expressed in categories

Name	Surname	Total	Phonics	Reading	Maths	Handwriting
ZZZ	YYY	2	1	2	4	4
ZZZ	YYY	3	3	2	4	3
ZZZ	YYY	3	2	3	4	0
ZZZ	YYY	3	3	2	4	4
ZZZ	YYY	2	2	1	2	4

Figure 6.7: Individual learner result tables – Prototype I

The section on exemplary learners and those in need of extra support was also maintained. A histogram of the learners' performance for each scale was however included to provide greater insight into the distribution of the learners' performance for the scale per school (see Figure 6.8).

5.4.3 EARLY READING SCALE

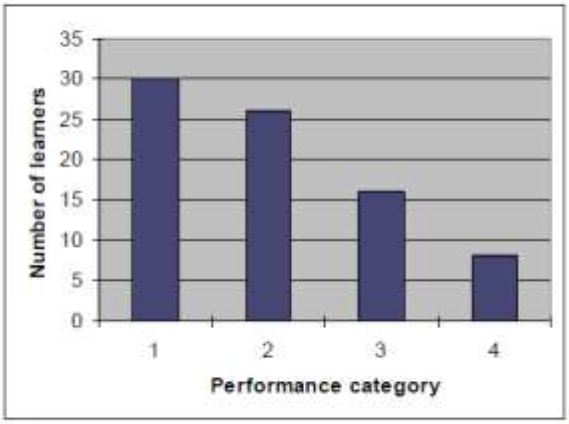


Figure 30: Distribution of learner performance according to performance category on the Follow-up Early Reading Scale

The majority of learners in your school had achieved a 1 (0-25%) or 2 (26-50%) on the Reading section of the assessment. The 30 learners who achieved a 1 have been identified as in need of additional support in this area.

Learners in need of support in the area of Early Reading:

Name	Surname
ZZZ	YYY
ZZZ	YYY
ZZZ	YYY
ZZZ	YYY
ZZZ	YYY

Figure 6.8: Histograms and identification of learners in need of additional support - Prototype I

A conclusion section was added to the report summarising the main strengths and weaknesses of the group tested for the school. An appendix was added with learner results presented as percentages for teachers who preferred working with percentages. The changes to the feedback session for Prototype I are now discussed.

6.1.1.2 Feedback Sessions

Representatives from schools were invited to attend and received the reports at the feedback session. With all the changes that had been implemented, the turnaround time for reporting had decreased by two weeks. The feedback session was shortened to 90 minutes and consisted of a presentation by the project leader, comprising:

- A shortened introduction to the CEA and the project (adapted)
- Description of the assessment (adapted)
- Presentation of the new report format (new)
- An example of how to make sense of the data in the new format (new)
- The future plans for the project
- An opportunity for discussion and questions

The focus shifted from introduction of the assessment to discussing how to understand the reports and make sense of the data. Schools were also provided with more opportunity for questions and discussion of application of results and problems in their schools. Schools were presented with a handout with the slides and an agenda to facilitate the process. The invitation to contact the research team for any further analysis or support was repeated officially. Refreshments were served and informal conversations took place between the research team and the schools' representatives about the project and individual school reports during the break.

6.1.2 Formative Evaluation of Prototype I

The formative evaluation of Prototype I took place through two processes. Firstly, school users were approached to evaluate the prototype and provide suggestions and priorities for improvement through the Delphi technique. Secondly, expert evaluators were asked to evaluate the prototype and provide verbal and written feedback. The research procedures for this formative evaluation are discussed below.

6.1.2.1 Selection of Participants

All schools in the SAMP sample participated in the ***feedback prototypes,*** i.e. all 22 schools in the sample that received reports were invited to the feedback session and received support material and telephonic support if required. The evaluation of Prototype I consisted of the Delphi technique, employed with the schools and an expert evaluator's examination of the reports and feedback session documentation. All schools were invited to participate in the Delphi technique and participation took place on a voluntary basis.

Sampling for Delphi technique

All schools were sent electronic facsimiles (faxes) of the forms to participate in the Delphi technique. Both expert and user evaluator were employed to evaluate Prototype I. All schools received all three rounds of documentation whether or not they had participated in the previous round. In some cases different schools responded to each of the three rounds of the Delphi technique. A non-response analysis was conducted (see Table 6.2). All English and Afrikaans schools participated in some round of the analysis and three of the seven Sepedi schools participated. In total, 17 of the 22 schools participated in the Delphi analysis. The low response rate for the Sepedi schools was attributed in part to problems with communication infrastructure at the schools and strike action at the time.

Table 6.2: Response analysis for Delphi technique –Prototype II

School Number	Language	Round 1	Round 2	Round 3
3	English			1
4	English	1	1	
6	Afrikaans			1
7	Sepedi	1		
8	Afrikaans	1		
10	Afrikaans			1
11	Sepedi		1	1
12	Afrikaans	1	1	1
16	English	1		1
17	English/Afrikaans	1	1	
18	English	1		
19	Afrikaans	1	1	1
20	English	1		1
21	English		1	
23	English	1	1	
24	Sepedi			1
25	Afrikaans	1	1	1
		11	**8**	**10**

Sampling of Expert Evaluators

Three expert evaluators and consultants were employed in this evaluation cycle. The three experts were linked to the CEA as consultants on a SANPAD collaboration project. The Dutch experts were therefore already involved in the value-added project and available to serve as expert evaluators. Each consultant's areas of expertise is summarised below:

- ***Expert 1:*** An expert in the field of education change, extensively involved in school-based research in South Africa. A qualitative research methodologist with specific expertise in computer-aided qualitative data analysis and a wide knowledge of the South African educational context.
- ***Expert 2:*** A Dutch academic involved in the development of the ZEBO school evaluation system in the Netherlands. A quantitative methodologist with expertise in Item Response Theory and multi-level modelling, and extensive knowledge of monitoring in the Netherlands education system.
- ***Expert 3:*** A Dutch academic with extensive national and international experience in educational evaluation and international schooling systems. Considerable expertise in the fields of design research and curriculum design.

The diversity of experience and expertise ensured that contextual knowledge of South Africa, monitoring knowledge, knowledge of design research methodology, qualitative and quantitative expertise were all presented by the expert evaluators.

6.1.2.2 Data Collection

This first method of choice for the evaluating of Prototype 1 was the Nominal group technique. Educators, HoDs and Principals from one school per language group were invited to participate. The Nominal group technique was originally chosen as the technique is often employed to generate a diversity of ideas with a level of anonymity, whilst still providing the opportunity of checking communication (Syque, 2007). The technique perfectly suited the role of the evaluators as revisors (Nieveen, 2009) in this phase. Unfortunately, during the first attempt at the Nominal groups, attendance was poor with only the selected English school participating and the participants tended to communicate naturally with each other, producing the 'group think' phenomena. This was contradictory

to the purpose of the approach to generate a multitude of ideas. Therefore, the approach was changed to the Delphi technique to address this issue.

The Delphi technique is similar to the Nominal group technique in that it is a group problem-solving and decision-making approach, but does not require face-to-face interaction (Michigan State University Extension, 1994). The technique starts by the posing of a specific problem to which participants anonymously contribute. This initial question statement and input is followed by a series of carefully designed questionnaires that incorporate summaries and comments from the previous rounds to generate and clarify ideas. The process concludes with a voting round in which the participants can indicate the priorities for the specific project (Dunham, 1995; Illinois institute of technology, nd; Williams & Webb, 1994). In this study, the questions were aimed at determining how the use of the feedback system for SAMP could be improved, with specific reference to three areas:

1. The feedback sessions at the University
2. The reports provided for each school
3. The support for understanding the information included in the reports

School representatives were also invited to comment on any other aspects of the programme.

Usually, e-mail communication is used in the Delphi technique, but many of the schools in the sample did not have access to the Internet so a slight adjustment was made to the technique with faxes being sent to and from schools. The technique proved more efficient and encouraged greater participation than the Nominal group technique with at least a third of schools in the sample contributing to each round of questioning. Diverse and rich ideas were also generated and discussed in relation to the feedback system.

The input from the three expert evaluators was less formal. The expert evaluators were provided with the reports and the support materials, as well as materials relating to the feedback session (including invites, agendas and slide presentations). The evaluators provided written and verbal comments and made recommendations including for improvements to the reports.

6.1.2.3 Data Capturing and Analysis

Data from the Delphi technique were captured electronically, both in textual forms for qualitative analysis and in *Excel* for descriptive analysis of frequencies. Feedback from the expert evaluators was captured through field notes from meetings, while notes on the reports were kept and captured electronically and electronic feedback received through e-mails was saved.

Data collection and analysis for the Delphi technique was iterative. The first rounds of data from the faxes were captured textually and analysed using *Atlas.ti* to generate themes of suggestions. The themes were constructed from the responses and presented in the second round of faxes. Further feedback and comments were requested in the second round and the themes revised, based on the feedback from round two, again employing *Atlas.ti*. The third round faxes summarised the comment and recommendation themes and allowed the schools to vote on the priority order for further development of the feedback system. These data were captured in *Excel* and descriptive statistics were generated to determine the priorities for further development of the feedback systems.

The data from the expert evaluators were thematically analysed to determine the focus for further development. These themes were compared to input from the users to determine which aspects should receive priority for further development. This process allowed for a grounding of the expert evaluations in the contextual needs of the users.

6.1.2.4 Results and Design Guidelines-Expert Evaluators

The evaluators commented on reports verbally and through written notes on the materials. The overall evaluation of the reports and feedback session was positive, the evaluators stating that the data were presented clearly and in an easily understandable format. A number of recommendations for improvement of the reports were provided:

1. ***Continued involvement of users*** in evaluating and improving the reports and feedback sessions

2. Ensuring of ***alignment of format and structure*** between the development of the secondary and primary school feedback systems

3. Dividing reports into a ***separate manual and data report***

4. ***Expanding the descriptions*** in the instrument manual with particular reference to the underlying skills involved in each subtest and the links to the curriculum

5. ***New instrument manual***

    - Including the ***reliability and validity*** data of the instrument development in the manual

    - Including a ***section on interpreting and using the data*** in the manual

    - ***Automating*** certain parts of the report generation to decrease turnaround time from assessment to reporting

6. ***Reports***

    - Changing the reporting to reflect learner scores in ***five instead of four categories,*** so that categories are less broad and allow for easier differentiation between learners

    - ***Comparing results to scores from previous years***. This would mean that reports could be generated, before all the schools were assessed, further decreasing turnaround time.

    - Placing ***standard error bars*** in the graphs to convey the message that there is uncertainty in the scores and a small difference in score should not be over-interpreted.

    - Ensuring that that the scales in the graphs on the ***Y-axis run from 0-100***.

The majority of these recommendations were only employed in Prototype III, to maintain consistency of reporting from the baseline to the follow-up assessment.

6.1.2.5 Results and Design Guidelines-Delphi Technique

Several themes emerged based on the Delphi technique, clustered according to material relating to the feedback presentations, reports and support at the University. The voting in the Delphi technique also resulted in priority values being assigned to the different suggestions for improvement.

I. FEEDBACK PRESENTATIONS

The input on the feedback presentations related mainly to the content, logistics and process of the feedback presentations.

I (a)Content

The schools generally expressed the view that the content of the feedback session was very informative and positive. Schools who have been involved in the project for longer indicated that the sessions seemed to be aimed at the newer schools and some of the information was repetitive. "*Session benefit only new schools who cannot interpret the results. If I know how to interpret the results, I do not need to attend session" (Q1-English-School 4).* Many of the schools indicated that they would appreciate more information on the questions used in the assessment and the link to the curriculum. "*Belangrik, sodat korrekte leeruitkomste aangespreek word aan die begin van die jaar"*[37] *(Q2-Afrikaans-School 12).* Some schools warned that while this would be good the information should be presented in such a way that it would not encourage teaching to the test *"…On other hand the test would not be successful as learners will be prepared" (Q2-English - School 23).*

I (b)Logistics

Some schools indicated that signage to the venue should be improved as the campus can be quite confusing. There was also a general request that the time of the feedback sessions should be brought forward slightly. The time of 14:00 was decided as most appropriate for the feedback sessions to allow for travel time. Both these suggestions were taken up without any further voting as they could easily be addressed. There were also suggestions to move the sessions to a

[37] English translation: *"Important so that the correct learning outcomes are addressed at the beginning of the year."*

more central location or conduct presentations at the schools individually. Some schools indicated a willingness to host the feedback sessions, but most schools stated that the University was a good central venue. Little support was shown for having individual feedback sessions at each school.

I (c) Process

Schools indicated that they were generally happy with the overall feedback process. One suggestion was that it might be beneficial to increase the participation by schools and allowing HoDs to make presentations. This suggestion was met with resistance from most schools who indicated that they were happy with the discussions and question-and-answer sessions already incorporated in the feedback: *"Nee, soos dit huidiglik is, is goed genoeg. Skole gee mos 'n inset gedurende die sessie, en kan vrae vra"*[38] *(Q2-School 19 - Afrikaans).* Some schools also indicated that asking HoDs to participate in such sessions would be unreasonable, given their current workload *"No time - too many obligations" (Q2-School 21 - English).* It was also suggested that the invitation to attend the feedback session be extended beyond just the principals, HoDs and Grade 1 educators to the whole Foundation Phase department. This idea was strongly supported by schools and it was further suggested that the Grade R educators attached to the schools could also benefit from attendance.

Priorities for improving feedback sessions

In order of importance, the following priorities for improvement were identified, based on the two rounds of questionnaires and voting of round 3:

[38] English translation: *"No, as it is currently is good enough. Schools give input during the session and may ask questions."*

Table 6.3: Delphi - Priorities for improving feedback sessions –Prototype I

Improvement Task	Mean Priority Level 1= Highest 4 = Lowest	Implemented
More information on the questions	1	Yes, included in manual Prototype III
Adjust the feedback sessions for old schools	1	Yes, format changed in Prototype II, feedback shortened and more of a focus on interpretation and new developments.
Include more people in the feedback sessions	2	Yes, all educators were always allowed to attend, invite reworded to indicate that all educators, Gr0-3 were welcome along with HoDs and Principals (Prototype III).
More central venue	3	No, not enough support
School presentations	4	No, not enough support
Improve the directions and signs to the venue.	No voting	Yes
Start feedback session earlier.	14:00 indicated as most appropriate	Yes, from Prototype II.

II. REPORTS

The feedback on the reports related mainly to their content, presentation and timing of the reports.

II (a) Content

Overall schools were happy with the content of the reports: *"Reports to schools were clear and covered all aspects of test" (Q1-School 16 - English)*. Many schools however suggested inclusion of additional variables in the report to expand the analysis and interpretation, e.g., age and demographics of learners, learners repeating Grade 1, home language and pre-school attendance: *"Yes, it makes a great difference in the results and performance" (Q2-School 23 - English).* It was also suggested that the reports include a breakdown of the learning outcomes. One school expressed that the reports seemed slightly large.

II (b) Presentation

The schools expressed the view that the reports were clearly written and understandable. *"The graphs give a clear indication of the results and we appreciate the information on who needs assistance and stimulation" (Q1-English School - 21). "All the information given was really simply done, understandable and user-friendly" (Q1-English-School 16).* Some schools indicated that sometimes the data were not a true reflection of certain learners'

capabilities. There was a request that schools should be grouped according to district for results: "*Areas will help to compare results and many areas work together" (Q2-School 23 - English)*. Some schools indicated that reporting should be done in all languages of assessment, though this was however not a high priority and schools noted that it might not be feasible. Some schools indicated that they were uncomfortable with comparison of results, even anonymously, with most indicating that the comparison was essential to compare school standards externally. *"Skole moet vergelyk word - Kan sodoende bepaal of skool hoë standaarde handhaaf"*[39] *(Q2-School12 - Afrikaans).* Schools indicated that it might be beneficial to have more than one copy of the reports provided to them; while others indicated that it was unnecessary. *"Think it is up to the school to photocopy the results to other teachers. This is a school management issue." (Q2-School 4 - English)*

II (c) Timing

Most of the schools indicated that they were happy with the current turnaround time. Some schools however noted that a shorter turnaround time would allow more time to work on identified problems.

Priorities for improving reports

The following priorities for improvement, in order of importance were identified, based on the two rounds of questionnaires and voting of round 3:

[39] English translation: *"Schools must be compared – Can then determine if school is maintaining high standards."*

Table 6.4: Delphi - Priorities for improving reports–Prototype I

Improvement Task	Mean Priority Level 1= Highest 5 = Lowest	Implemented
Report on additional variables	1	No, attempted to include, but schools failed to provide additional data
Report on overall trends across schools	2	Yes, reported in feedback sessions
Group schools - district or area	3	No, only reported per language
Provide reports earlier	4	Yes, report automation from baseline 2009.
Minimise comparison with other schools	4	Yes, from 2009 main comparison to average and own school's previous performance.
Produce reports in Afrikaans and English	4	No. Would require that all testing languages be used for reporting, not feasible.
Clearly state the learning outcomes	4	Expanded and stated in manuals after separation, baseline 2009.
Provide 2 copies of reports	5	No, additional cost not justified, responsibility of school.

III. SUPPORT

Feedback on the support provided related mainly to the administrative aspects, professional development, materials for intervention and opportunities to network with other schools.

III (a) Administrative

It was suggested that informed consent letters be provided to parents at enrolment the year before the assessments. This enjoyed support from the schools, who indicated that it would also help schools to plan their academic calendars with the dates formalised. *"This can be a proactive move from your side and school's side. When planning is made at school level, your evaluation dates will be properly planned for" (Q2-School 4 - English).*

III (b) Professional development

Schools were in favour of expanding the professional development activities and support beyond only Grade 1 educators, specifically to include Grade R/0 educators (Q1-School7 - Sepedi). *"Belangrik. Sal Gr R leerders meer toerus vir Graad 1 – Dat hulle skoolgereed is"*[40] *(Q2-School12 - Afrikaans)*. Some schools indicated that support for development and improvement of Grade R/0 would be appreciated as they identify areas for improvement here and evaluate progress

[40] English translation: *"Important. Will better prepare Grade R learners for Grade 1 – So that they are school-ready."*

for these teachers too (Q2-School 4 - English). Educators also indicated that it would be meaningful for them to attend the assessments, to have a better idea of content, the procedures and how the children react to the fieldworkers in the testing situation (Q2-School 23 - English).

III (c) Intervention materials

Schools asked for the provision of intervention materials to help support the addressing of areas of difficulty that had been identified. *"Teachers need support materials that you are using" (Q1-School 7 - Sepedi).*

III (d) Opportunities for networking with other schools

Schools indicated that they enjoyed being able to interact with other schools to address issues. They found these opportunities at the feedback session beneficial and asked for more such opportunities (N-School 23 - English).

III (e) Priorities for improving support

The following priorities for improvement, in order of importance, were identified based on the two rounds of questionnaires and voting of round 3:

Table 6.5: Delphi - Priorities for improving support–Prototype I

Improvement Task	Mean Priority Level 1= Highest 4 = Lowest	Implemented
Provide support to Grade R/0 Educators	1	Yes, included in feedback sessions, electronic resource Prototype III
Consent letters available during enrolment	2	Yes, baseline 2009 letters provided to schools in latter half of 2008 Prototype III
Allow educators to observe assessments	3	Yes, always been allowed, made schools aware at feedback that they may observe Prototype II
Workshops for teachers - project and application	4	No, only individual sessions and meetings with schools as requested. Schools made aware of this option again in Prototype II

IV. MONITORING SYSTEM ITSELF

Although schools were asked only to provide ideas around the reports, feedback sessions and support for interpretation and implementation, they also provided some additional information on the project for improvement. These ideas were noted and implemented where possible, but not circulated for voting or further discussion, as it was not the current focus of development. The main input around

the assessments was a request for an expansion of the sample within the participating schools:

- to include all grade 1 learners: *"Is it possible to test all the Grade One's +/- 200?" (Q 1-School 8 - Afrikaans)*
- to include other grades: *"How about evaluating grade 4s.(sic)"(Q1-School 7 - Sepedi)*
- by increasing the number of schools in the sample
- by increasing the frequency of the monitoring to more than twice a year

This input is encouraging as it indicates that the monitoring and feedback was valued and that expansion of the project would be welcomed. Schools also provided some input on further fieldworker training.

The evaluation data on the feedback sessions and support were incorporated into the feedback session and support for Prototype II. The reports however maintained their structure from the baseline to ensure report consistency for users from the baseline to the follow-up assessment. Only minor changes were performed on the report formats for Prototype II. Guidelines that had a large impact on the reporting format were only incorporated in Prototype III – Baseline 2009.

6.1.3 Cycle 2 (Prototype II, Follow-up 2008)

Feedback prototype II was developed based on the previous evaluations, with adjustments made to the feedback sessions, communications and logistics. Only minor changes were made to the reports, with major structural changes to them only implemented in Prototype III. This staggering of changes was essential to ensure that the baseline and follow-up reports for 2008 remained similar in structure and format so that changes would not interfere with interpretation and comparison of the baseline and follow-up results for 2008. The formative evaluation of Prototype II was more directly focused on the feedback session and employed school-users as evaluators. The guiding research and evaluation question for this cycle was therefore still question 3, but with a focus on the feedback sessions:

Research question:

3. *What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?*

The ***design guidelines*** generated related specifications to establish use of the feedback system, specifically the feedback sessions:

2. *Establishing conditions for use:* This development stage was aimed at improving the feedback session component of the feedback system. For this cycle the focus was on examining the design of the feedback sessions in detail.

The evaluation in this cycle focused specifically on the ***evaluative foci*** of relevance, consistency and expected practicality, with specific reference to the feedback sessions and reports. In the following section, Prototype II of the feedback system, developed and implemented based on the guidelines of Cycle 1 is introduced.

6.1.4 Prototype II – Follow-up 2008

The same sample of 22 schools from the baseline 2008 assessment participated in the follow-up assessment. Therefore, 1,390 learners from the baseline assessment were assessed in the follow-up for 2008. There was a drop out rate of 9.4%, due largely to learner absenteeism and migration. All the schools also participated in the feedback system.

All the schools in the sample received paper-based reports that included an instrument manual section describing the different subtests and scales. Teachers, principals and HoDs were invited to attend the feedback session, while schools were also invited to contact the CEA if there were any questions or if there was any help or support needed with interpretation or the results and planning. Additionally brochures for parents on providing literacy support and printed reading support materials were provided to schools for their own use at the feedback session, as indicated by the requests for additional support in Cycle 1. The components of Prototype II are shown in Table 6.6.

Table 6.6: Prototype II – Follow-up 2008 components

Component	Description
Paper-based follow-up report	Produced for each school individually and handed out at the feedback session
Follow-up feedback session	Principals and teachers from all participating schools invited to the University of Pretoria
Literacy support materials	Some resources printed and handed out to schools at the feedback session
Telephonic, written and face-to-face communication	On an ad hoc basis as required

The changes to the paper-based reports and feedback sessions are discussed below.

6.1.4.1 Reports

The overall structure of reports remained consistent from the baseline to the follow-up report. The reporting of follow-up scores however presented the gains data differently than the pre-existing system, in that the difference between the baseline and follow-up score was shown as a gain or loss on the same bar (see Figure 6.9)

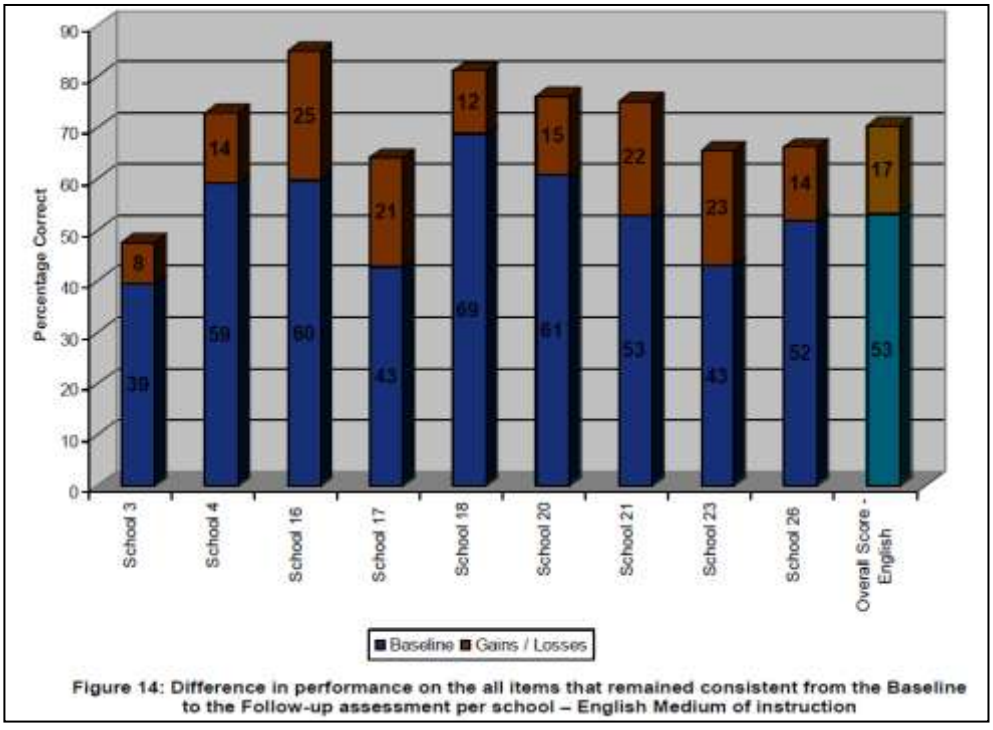


Figure 6.9: Overall school performance in terms of gains and losses in percentage correct – Prototype II

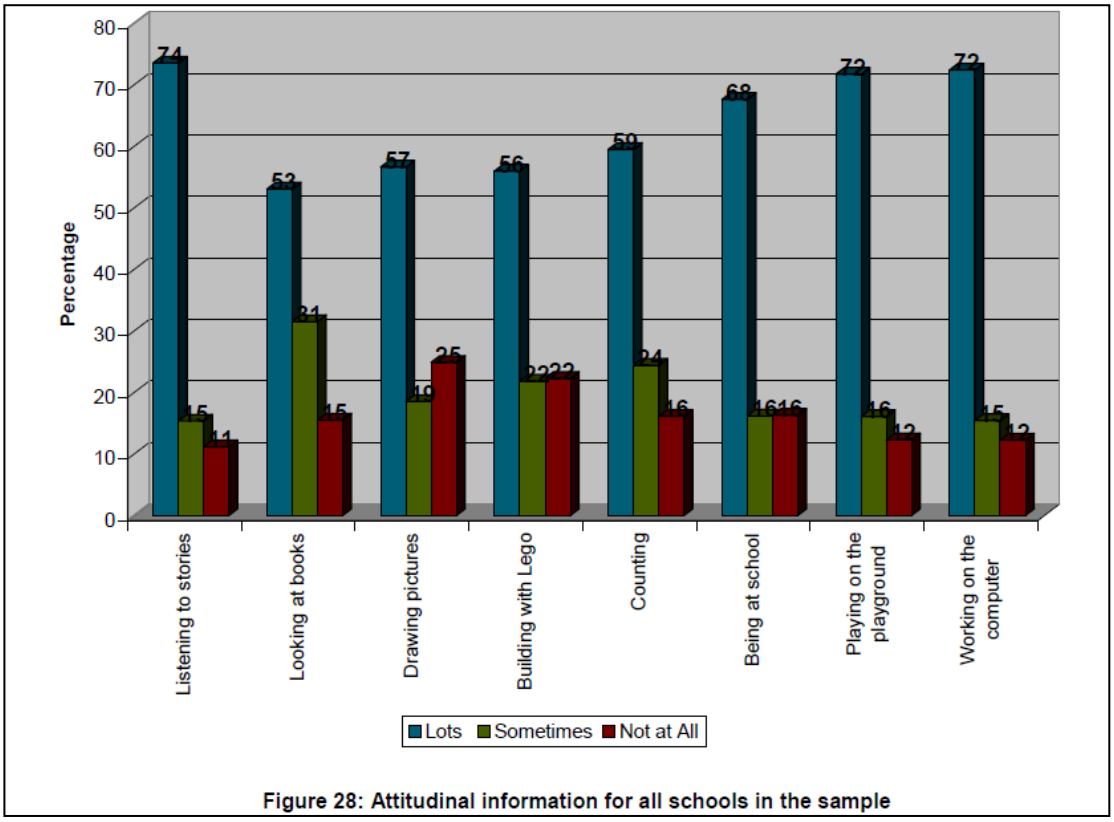


Figure 6.10: Attitudinal information for learners on school-level – Prototype II

Additional information was also included in the form of attitudinal data for the learners in the school (see Figure 6.10, above). The attitudinal data were first presented as an average for all the schools across the sample and then for the specific school.

The changes for this prototype focused on the feedback session. The feedback sessions for Prototype II is discussed below.

6.1.4.2 Feedback Sessions

School representatives were invited to attend the feedback session, where they received the report. As per the evaluation of Cycle 1, care was taken to ensure that schools knew they could bring a group of people, including Principals, HoDs and teachers from Grade 1 and other related years. The time of the feedback session was moved to 14:00, the time agreed on as most convenient for all parties in the Delphi technique and the placement and number of directions to the venue was increased. It was also made clear in the feedback session that teachers were welcome to observe assessments and speak to fieldworkers. The

feedback sessions was kept to 90 minutes and consisted of a presentation by the project leader pertaining to:

- Introducing the CEA and the project
- A shortened description of the assessment
- An example of how to interpret the data (new)
- An example and discussion of possible actions based on the interpreted data (new)
- The future plans for the project
- An opportunity for discussion and questions

The focus shifted from introduction of the assessment to discussion of how to understand the data and use it for planning and action in the school. Schools were also given an opportunity to work through an example and provide their own interpretation. Schools were presented with a handout with the slides and an agenda to accompany the presentation. The invitation to contact the research team for any further analysis or support was extended. Refreshments were also served and informal conversations took place between the research team and schools about the project and individual school reports during the break.

6.1.5 Formative Evaluation of Prototype II

The formative evaluation of Prototype II focused on the feedback session. Schools were asked to evaluate the feedback sessions by completing an extensive questionnaire. The research procedures for the formative evaluation of Prototype II are discussed below.

6.1.5.1 Sampling

All schools were invited to the feedback session for the follow-up assessment 2008. As the questionnaires focused on evaluating the feedback session, only schools who attended the follow-up SAMP 2008 feedback session were provided with them. Eighteen responses from 15 different schools were received from teachers, HoDs and principals who attended the session (see Table 6.7)

Table 6.7: Number of respondents for feedback questionnaire –Prototype II

School name	Grades Represented	Language of instructions	Number or questionnaires
School 4	1	English	3
School 5	1	Sepedi	1
School 6	1	Afrikaans	1
School 7	2	Sepedi	1
School 8	2	Afrikaans	1
School 11	1	Sepedi	1
School 16	1	English	1
School 18	1	English	1
School 20	1	English	1
School 21	1	English	1
School 23	1	English	1
School 24	1	Sepedi	1
School 25	1	Afrikaans/English	1
School 26	1	English	1
School 26	1,3	English	2
Total			**18**

6.1.5.2 Data Collection

Data collection to inform the development of Prototype III consisted of questionnaires specifically designed to evaluate the feedback sessions of Prototype II. The questionnaires were administered to teachers, HoDs and principals from schools in the feedback system sample. The questionnaire (see the audit trail DVD) focused on the feedback sessions and was administered to persons who attended the follow-up 2008 SAMP feedback session. The questionnaire was aimed at gaining a deeper understanding of the schools perception of the feedback sessions and to elicit suggestions for further improvement. It was important to establish if the sessions created the appropriate conditions for use of the SAMP monitoring system and data.

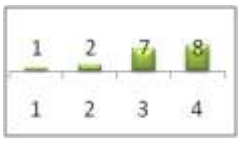
6.1.5.3 Data Capturing

The questionnaire data were captured in *Excel* spreadsheets. Open comments by the respondents were also captured electronically, to be analysed thematically, but the quantitative component was the focus of the evaluation for this cycle.

6.1.5.4 Data Analysis

Descriptive statistics were employed to analyse the questionnaire data using *Excel* and *SPSS*. The comments made on the questionnaire were also analysed thematically to provide a richer context for the quantitative data.

6.1.5.5 Results and Design Guidelines

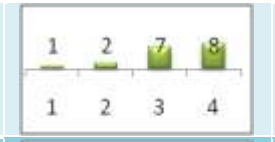
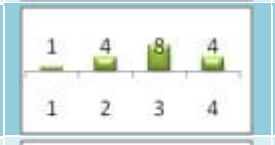
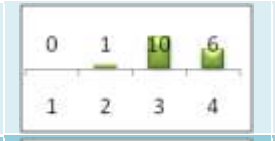
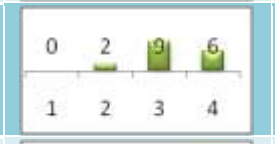
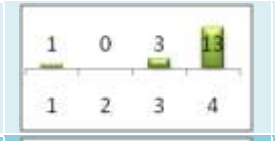
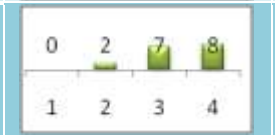


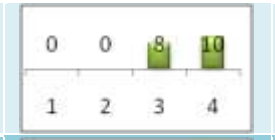
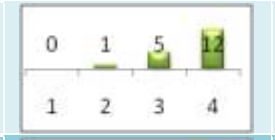
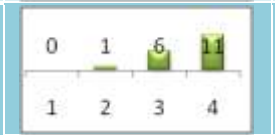
As such a large amount of data were collected through the questionnaires, graphs representing the frequency of responses for each question were employed to summarise and condense the results The numbers below the X-axis: 1,2,3,4 represents the different categories on the scale, 1=poor to 4=Excellent. The numbers above the X-axis represent the frequency of responses for each category. In this example, eight respondents rated this aspect as 4, or Excellent. This type of frequency distribution representation is used throughout the thesis. Please also note that while N=18 for this questionnaire, not all respondents answered all questions resulting in a lower response rate for some questions.

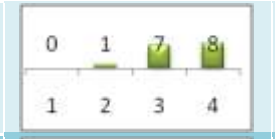
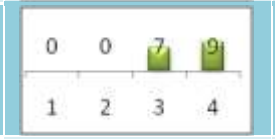
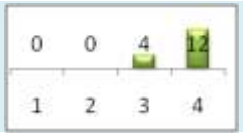
The results of the feedback questionnaire are summarised in Table 6.8 Overall, the level of satisfaction with the feedback session was very high, with an average rating of 3.5 out of a possible 4. The lowest ratings were still around logistical matters, schools indicating that the time of the sessions was still difficult, but most commenting that there could be no better time as it had to be in the afternoons.

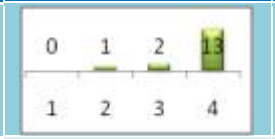
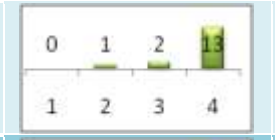
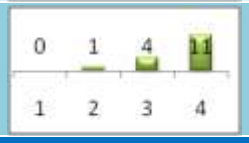
Schools noted that the feedback sessions were a priority for them and they would “make time”. Some noted that they were still having trouble with the directions to the venue. The feedback sessions were rated very highly (mean score of 3.8) in terms of being understandable, addressing school and teacher concerns and helping to make sense of the data.

Table 6.8: Cycle 3 - Summary of feedback questionnaire results

Question	Average rating 1=Poor - 4=*Excel*lent	Frequency distribution for response	Selected comments
1. Directions to the venue	3.2		Got lost on the last turn just after the second staircase.
2. The agenda of the session	3.4		I like the short and sweet version.
3. The length of the feedback session	3.4		-
4. Scheduling of the feedback session	2.9		There can't be a good time ... This is important to be done and time made for the info. Thank you.
5. Timeliness of the information	3.3		Good, we can now judge whether our retentions and progressions are on track. Information very good but timing inconvenient as we are really busy with reports and end of year work.
6. Relevance to my concerns	3.2		All learners should be assessed.
7. Opportunities to ask questions	3.7		-
8. Interaction between participants and the presenter	3.4		Your team are also very friendly when at our school. They are also accommodating. Thank You.

Question	Average rating 1=Poor - 4=*Excel*lent	Frequency distribution for response	Selected comments
9. Knowledge of the presenter	3.6		Liz is always friendly, efficient and very professional.
10. Handouts and resource package provided	3.7		*Excel*lent, we are glad to track our learners' progress.
11. Quality of the presentation	3.6		Very professional and well done.
12. The environment in which the feedback is given	3.6		Become too small (Maybe everyone didn't reply & your room wasn't large enough?)

Question	Extent : 1=Not at all - 4=Completely	Frequency distribution for response	Comments
13. Discussion of further developments	3.4		Money is one difficult concept for our learners. Mathematics is becoming easier… the problem we encounter is phonic awareness and word recognition
14. Did the feedback session meet your expectations?	3.6		-
15. Did the information provided assist with the understanding of the report content?	3.8		-

Question	Extent : 1=Not at all - 4=Completely	Frequency distribution for response	Comments
16. Was the presenter helpful in addressing concerns raised?	3.8		Very ☺
17. Was the presentation clear and understandable?	3.8		-
18. Did the feedback session provide the opportunity to learn something useful?	3.6		-

Question	Selected Themes
19. What were the strengths of the feedback session?	**Overall:** Provided insight on the performance of learners at school compared to other schools. | Will assist in changing focus on aspects that needs attention next year. | Keep up the *Excel*lent work!
	Presentation style: *Baie deeglik, Baie uitnodigend aangebied.*[41] |…Presenter is dynamic and well prepared.
	Anonymity: Reading results for each school without naming schools.
	Interpretation and analysis support Analysation (sic) of problematic topics in different learning area. | A lot of work that I did not understand is now clear… | Interpretation of feedback was simple to understand...
	Manual and report: *Excel*lent study guide. The feedback book is *Excel*lent, well put together…
20. What are perceived weaknesses of the feedback session?	**Timing of the session:** 1 December! Maybe earlier
21. Do you have any other comments regarding the feedback session, which has not been addressed?	**Overall:** I enjoyed the feedback. It was fruitful and I hope to implement it in my class.
	Support: Assisting learners with critical thinking skills in mathematics.
	Fieldworkers: Ladies are recommended to do this assessment as lower grade learners are taught by ladies. They might be frightened to be tested by men.

[41] English translation: *"Very thorough. Presented in a very inviting manner."*

In the general comment section, the schools noted the following ***strengths***:

- ***Overall:*** The feedback provides insight into comparative school performance and what the focus should be for the following year.
- ***Presentation style:*** The presentations were thorough and inviting. The presenter was clearly well prepared, inviting and professional.
- ***Anonymity:*** Comparative results were provided without naming schools
- ***Interpretation and analysis support:*** The presentation helped to make sense of problematic topic areas, and helped in understanding what was unclear. The interpretation and feedback were perceived as clear and simple.
- ***Manual and report:*** Well presented and structured.

The one ***weakness*** that was noted related to the timing of the feedback sessions and it was suggested that they take place earlier. A request was made for support in addressing areas of concern, such as critical thinking in Mathematics.

The evaluation indicated a high level of satisfaction with the feedback sessions and even the addition of the small amount of support material was appreciated. It seemed that the feedback session structure and process were effective and the focus could now shift to examining the reports and how to facilitate the use of the feedback system in the various contexts.

Design guidelines from Feedback Evaluation Questionnaire

1. ***Improved turnaround time is important*** in increasing the relevance, usefulness and efficacy of the feedback. This is not only important so far as the reports are concerned, but also for the scheduling of the feedback sessions.
2. ***Feedback must be linked with resources and suggestions for action:*** the presence of links to support material facilitates the use of

feedback. Not having access to resources or being aware of resources to address issues raised in feedback often hinders use of feedback. Providing such links facilitates the use of the feedback

3. ***Opportunity for two-way communication*** is important. Such opportunity can be created through formal discussions during feedback. An opportunity for informal one-to-one discussions with school representatives during for instance refreshments is important to address school specific concerns in a confidential manner. The atmosphere created during feedback should be ***non-judgemental, constructive and invite participation***. Feedback must also be ***clear, concise and simple,*** so conversation can focus on interpretation and application, not only understanding of the data.

6.2 Conclusion

In this chapter the first two cycles of the Prototyping Phase were documented with emphasis on establishing the conditions for use of the feedback system. This was achieved by examining and improving the different components of the feedback session. The design guidelines from these cycles informed the development of the third prototype. Prototype III, as part of Cycle 3 is discussed in Chapter 7. The next chapter focuses on transforming conditions of use into action and planning in schools.

CHAPTER SEVEN
Prototyping Phase: Transforming conditions for use into use (Cycle 3)

> We are concerned that schools are being pushed and enslaved by data rather than being steered by leaders, with data providing information that they can use to engage in thoughtful planning and make reasoned and targeted decisions to move towards continuous improvement (Earl & Katz, 2006, p. xiii)

The first part (two cycles) of the Prototyping Phase focused on establishing the conditions for use of the feedback system (See Chapter 6). This chapter is the natural progression of the Prototyping Phase to examine how to transform these conditions into use by schools. It takes a closer look at the research design and results for the third of the three design cycles employed during the Prototyping Phase.

The complete research cycle discussed in this chapter consisted of the development of successive prototypes of the feedback system, which were formatively evaluated to inform the development of the final prototype for this thesis. The focus of this chapter is illustrated graphically in Figure 7.1. The pre-existing system and feedback prototypes are illustrated in blue and the evaluation activities in green. Every full design cycle consists of the prototype adaptation followed by implementation and corresponding formative evaluation of that prototype.

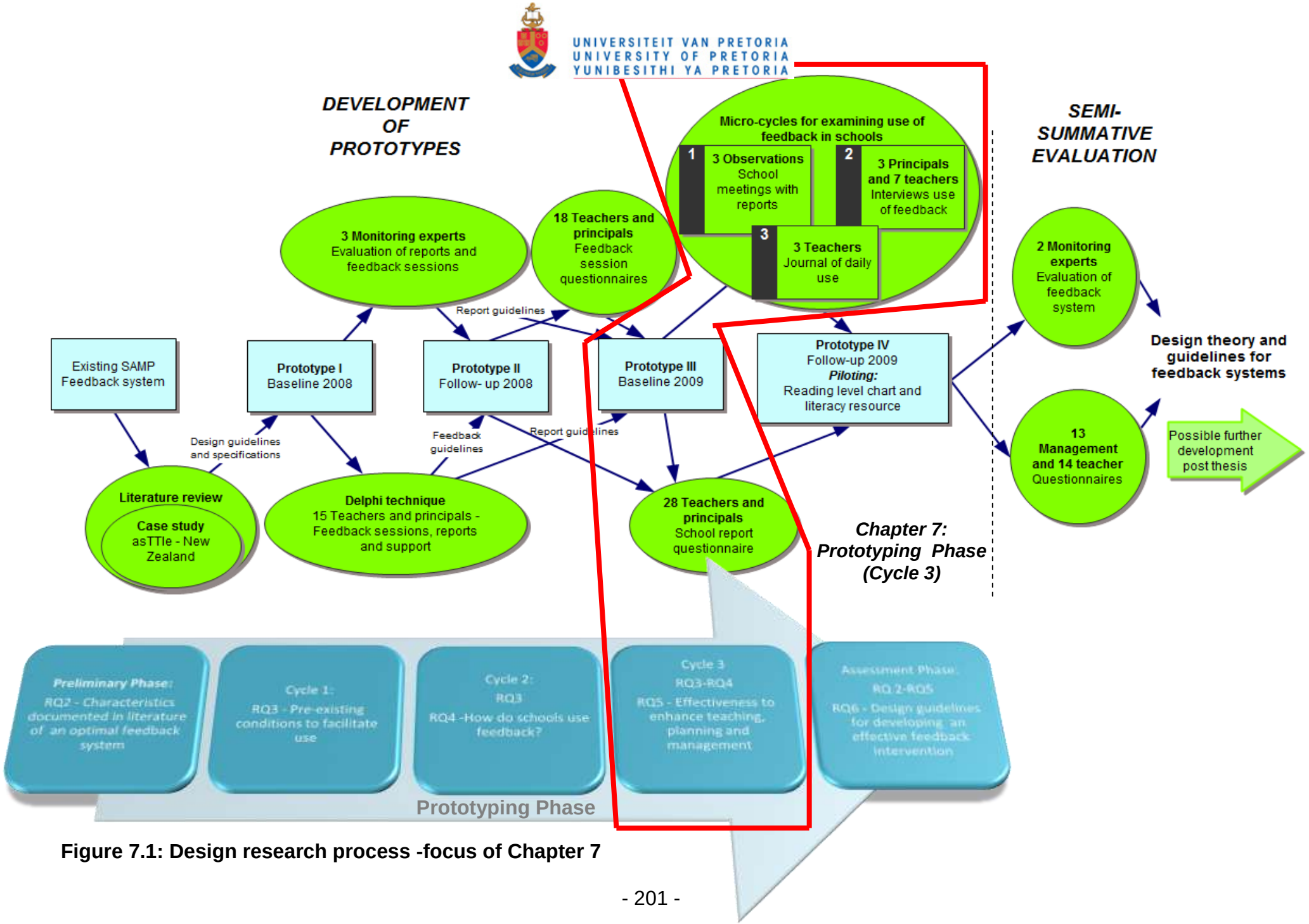


Figure 7.1: Design research process -focus of Chapter 7

In the following section the third cycle of the Prototyping Phase is discussed, beginning with a description of the prototype, followed by a discussion of the research design used to evaluate it. The evaluation activities are guided by the research questions being addressed and the evaluative focus for the specific cycle. While there is a specific focus for each cycle there is always some deliberate overlap and a cycle may address some aspects of other research questions or evaluative foci. The research procedures for each cycle examine the sampling, data collection and instruments, analysis and discussion.

7.1 Cycle 3 (Prototype III - Baseline 2009)

Feedback Prototype III incorporated the learning from Cycle 2 as well as the larger structural changes to the reports in order to transform the conditions for use of the feedback into action in the schools. The evaluative focus for this cycle was twofold:

1. To evaluate the functioning of the new report format, to facilitate further adaptation.
2. To evaluate the functioning of the complete learner performance feedback system, as a facilitator of evidence-based practise and planning in the schools.

The formative evaluation for this cycle employed a questionnaire to evaluate the functioning of the new report format. The transforming of the feedback into action in schools was investigated through observations of school meetings, structured reflective journals kept by schools and group interviews with teachers, HoDs and principals to follow up on the observations and journal data. The focus was thus to determine how schools were interacting with the feedback and the barriers or facilitating factors to employing it in each school's specific context. Questions aimed to establish how these aspects were being used in the context for which they were designed and to inform further improvement of the intervention. Expressed or presented use, along with limited data on actual use were included for this thesis.

Cycle 3 therefore focused on the following ***research questions***:

3. *What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?* ***(limited to the reports)***
4. *How do schools use feedback?*
5. *How effective is the feedback system in enhancing classroom practices, management and planning activities?* ***(focussing on expected efficacy)***

The evaluation information served to provide ***design guidelines*** relating to the report. Design guidelines to facilitate transforming the feedback into action were also developed:

3. *Establishing conditions for use:* This development stage aimed at improving the components of the feedback system. For this cycle there was a detailed examination of the design of the reports.
4. *Transforming conditions for use into action in schools:* This development stage followed school processes in employing the feedback system on a management, planning and classroom practice level. The focus was on the application of the report and feedback sessions data in the schools, along with the use of support materials, instrument manuals and linked intervention materials. This exploration was the focus of the micro-cycles of evaluation of Prototype III.

The evaluation in this cycle focused specifically on the ***evaluative foci*** of actual practicality and expected efficacy with specific reference to the reports and manner in which feedback is used in schools. In the following section Prototype III of the feedback system developed and implemented in Cycle 3 is introduced shortly.

7.1.1 Prototype III – Baseline 2009

A sample of 22 schools (Afrikaans, English and Sepedi) was maintained for the baseline 2009 assessment, with exactly 1,700 Grade 1 learners assessed. All the schools also participated in the feedback system and all received paper-based reports, but this time delivered prior to the feedback session. The report was also

split into two parts to create a report and a separate instrument manual. Teachers, principals and HoDs were invited to attend the feedback session. Schools were also invited to contact the CEA if there were any questions or if there was any help and support needed with interpretation and planning. At the feedback session each school received a DVD or CD containing freely available literacy, phonics, numeracy and mathematics materials. The components of Prototype III are shown in Table 7.1.

Table 7.1: Prototype III – Baseline 2009 components

Component	Description
Paper-based baseline report	Produced for each school individually and delivered to each school a week before the feedback session
Paper-based instrument manual	Delivered to schools with reports to accompany report
Baseline feedback session	Principals and teachers from all participating schools invited to the University of Pretoria
Electronic resource	Electronic collection of literacy, phonics, numeracy and mathematics materials produced on CD/DVD and provided to each school at the feedback session. Including an electronic copy of the report (PDF) with a learner dataset for the school in Excel
Telephonic, written and face-to-face communication	On an ad hoc basis as required

The changes to the paper-based reports, manuals and feedback sessions, as well as the newly included electronic resources are discussed below.

7.1.1.1 Reports

The major changes to the reports noted from Cycle 1 were implemented in this cycle. The first involved splitting the report into a separate report and instrument manual to facilitate interpretation and shorten it. The adapted structure of the reports was as follows:

1. Introduction to the SAMP project
2. Results per school ***(adapted)***
    a. Individual results
    b. Notes on interpretation

c. Learner results table

d. Learners at risk

e. Exceptional learners

3. Conclusion and recommendations ***(adapted)***

Section 1, the introduction to the project, remained similar to that used in Prototype II. A number of changes were implemented in the report production, data representation and report delivery to improve the turnaround time and ensure that the schools could receive reports prior to attending the feedback session. Firstly, data were no longer reported as comparative data for all schools for the current year. School data for a particular school were represented, compared to the school's performance in the previous year as well as other schools' results in the previous year (see Figure 7.2). This shift had the benefits of allowing report writing to commence earlier, as all the fieldwork did not need to be completed and the reporting style also allowed schools to view results more easily in the context of the previous year's performance.

Table 1: Average percentage per scale and total for Afrikaans learners

Scales	Results from 2008 SAMP baseline assessment							School 12 - 2009
	School 6	School 8	School 10	School 12	School 19	School 25	Overall Score - Afrikaans	
Handwriting	86	83	85	86	82	80	**84**	78
Early Phonics	67	65	51	56	71	55	**62**	56
Early Reading	46	50	38	42	45	31	**43**	41
Early Mathematics	84	86	78	85	87	78	**83**	84
English Additional	72	71	61	58	74	58	**66**	63
SAMP Total	**66**	**67**	**56**	**61**	**68**	**54**	**63**	**60**

Figure 7.2: Example of comparative results per language group per scale - comparison of 2009 school results to 2008-2009 – Prototype III

Secondly, the report writing process was automated, but due to financial limitations it was accomplished by use of advanced report writing techniques incorporating the *Microsoft Office Suite*. All programming was done by the researcher. This limited cost and ensured that any technical errors experienced during the design and development process could be addressed in-house.

Zapawa (2005, pp. 5-6) notes a number of other benefits to using the *Microsoft Office Suite* for advanced report-writing over custom purchased and designed systems, including:

- Real-time retrieval of data across the suite
- Sorts, breaks and totals can easily be removed and modified
- Some report types are dynamic and powerful
- Reports are cost effective
- Report production time is often shorter than with competing software
- Reports are integrated with other Microsoft Office Suite program[me]s

This report automation reduced turnaround time, decreased opportunities for transposition errors and therefore reduced costs related to manual report production[42]. As *Microsoft Office* was used, it also meant that datasets could easily be provided to schools on the electronic resource CD/DVD in a format that could be accessed through *Excel* for further analysis. This automation process was in line with the findings from the case study and literature review that the use of ICT was recommended for sustainability and expansion.

The automation process also influenced the representation style of the graphs, reducing the numbers of bars to three. Only the school average for 2009 and 2008 as well as the language group average for 2008 was reproduced, as opposed to the individual results for each school (see Figure 7.3). The presentation style was updated with the *2007 Office* package to give a more professional appearance. The vertical axis was also pre-set to a scale from 0-100 to ensure consistency throughout.

[42] Although the report process was automated, interpretation and discussions still had to be done individually to ensure that the quality of data representation was not negatively impacted.

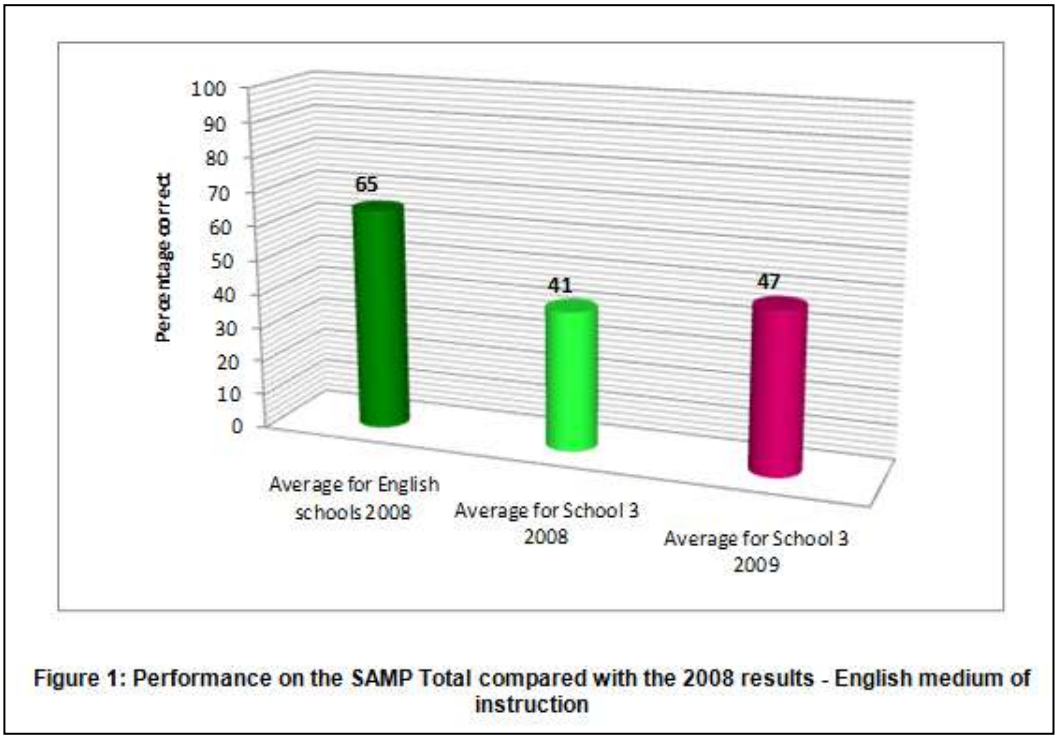
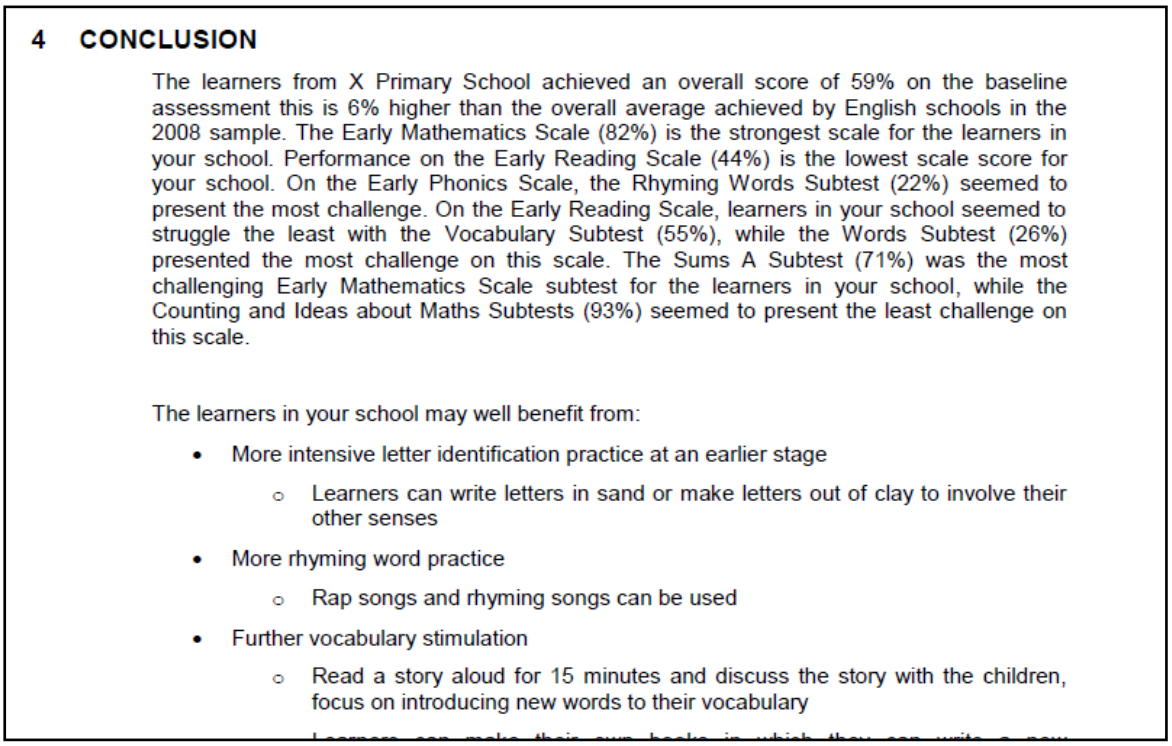


Figure 7.3: Example of overall performance graph for school 3 compared to English medium schools – Prototype III

Thirdly, reports were no longer handed out during the feedback session. A team member would visit each school and deliver the report and manual to the school a week prior to the feedback session.

The report conclusion section was expanded and the discussion and summary of results supplemented with a list of suggested activities tailored to the results of the specific school. A comparative component was also added to this section, discussing changes in the school results over a two year period (see Figure 7.4).



4 CONCLUSION

The learners from X Primary School achieved an overall score of 59% on the baseline assessment this is 6% higher than the overall average achieved by English schools in the 2008 sample. The Early Mathematics Scale (82%) is the strongest scale for the learners in your school. Performance on the Early Reading Scale (44%) is the lowest scale score for your school. On the Early Phonics Scale, the Rhyming Words Subtest (22%) seemed to present the most challenge. On the Early Reading Scale, learners in your school seemed to struggle the least with the Vocabulary Subtest (55%), while the Words Subtest (26%) presented the most challenge on this scale. The Sums A Subtest (71%) was the most challenging Early Mathematics Scale subtest for the learners in your school, while the Counting and Ideas about Maths Subtests (93%) seemed to present the least challenge on this scale.

The learners in your school may well benefit from:

- More intensive letter identification practice at an earlier stage
  - Learners can write letters in sand or make letters out of clay to involve their other senses
- More rhyming word practice
  - Rap songs and rhyming songs can be used
- Further vocabulary stimulation
  - Read a story aloud for 15 minutes and discuss the story with the children, focus on introducing new words to their vocabulary

Figure 7.4: Expanded Conclusion Section – Prototype III

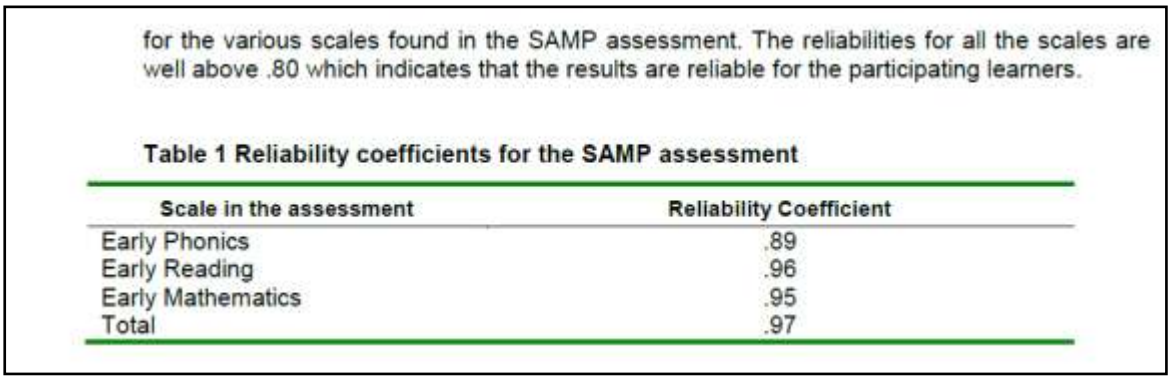
In addition to the paper-based report, the report was also saved as a PDF document. The electronic version of the report was saved on the electronic resource for each school to allow them to print out more full-colour reports if required. Schools could also use the report in electronic format.

The instrument manuals that were separated from the report are now discussed below.

7.1.1.2 Instrument Manuals

All the schools in the sample received a separate instrument manual delivered with the reports. The manual still contained a description of the project and the various subtests and scales in the assessment. The description of the link between the curriculum and each subtest, as well as the underlying skills assessed in each one was expanded to facilitate use of the data.

The new manuals were expanded to include sections on the quality of the data with reference to validity and reliability (see Figure 7.5) as well as a section on interpretation (see Figure 7.6). The interpretation section provided guidelines on how to interpret and use each section of the report in detail.

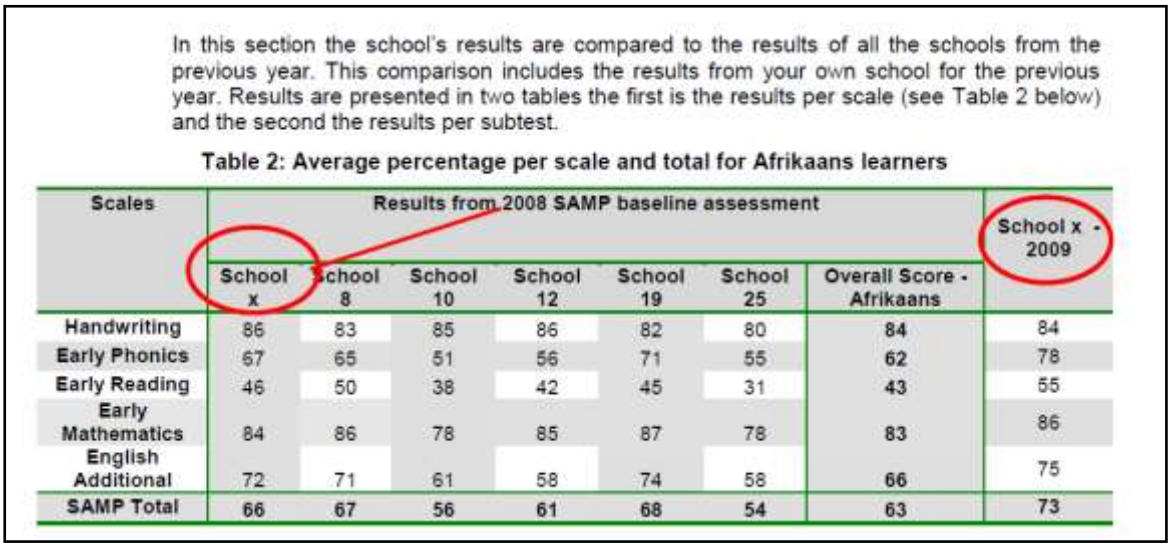


for the various scales found in the SAMP assessment. The reliabilities for all the scales are well above .80 which indicates that the results are reliable for the participating learners.

Table 1 Reliability coefficients for the SAMP assessment

Scale in the assessment	Reliability Coefficient
Early Phonics	.89
Early Reading	.96
Early Mathematics	.95
Total	.97

Figure 7.5: Manual- Excerpt from the Quality of Data Section – Prototype III



In this section the school's results are compared to the results of all the schools from the previous year. This comparison includes the results from your own school for the previous year. Results are presented in two tables the first is the results per scale (see Table 2 below) and the second the results per subtest.

Table 2: Average percentage per scale and total for Afrikaans learners

Scales	Results from 2008 SAMP baseline assessment							School x - 2009
	School x	School 8	School 10	School 12	School 19	School 25	Overall Score - Afrikaans	
Handwriting	86	83	85	86	82	80	84	84
Early Phonics	67	65	51	56	71	55	62	78
Early Reading	46	50	38	42	45	31	43	55
Early Mathematics	84	86	78	85	87	78	83	86
English Additional	72	71	61	58	74	58	66	75
SAMP Total	66	67	56	61	68	54	63	73

Figure 7.6: Manual- Excerpt from the Notes on Interpretation Section – Prototype III

The feedback session of Prototype III is discussed below.

7.1.1.3 Feedback Session

All schools were invited to attend the feedback session. The session was different from the previous sessions, as schools had already received their reports. Therefore, no reports were handed out at the session, though each school received a data reference sheet for 2009, showing the school comparative data for 2009 that was not covered in the report. Each school also received their individual electronic resource. The feedback sessions was shortened to one hour and consisted of a presentation by the project leader followed by an introduction to the electronic resource. The feedback session pertained to:

- A shortened introduction to the CEA and the project ***(modified)***
- An overall description of the assessment

- Presentation of additional data on the constitution of the sample, based on data provided by the schools ***(new)***
- Presentation of the new report format ***(new)***
- Introduction of the new data reference sheets handed out at the feedback session ***(new)***
- An example of how to make sense of the data in the new format
- Introduction of the electronic resource ***(new)***
- An opportunity for discussion and questions

The feedback session still followed the same order, with handouts of the presentation being presented and refreshments being served. The focus for this feedback session was on familiarising the schools with the new report format, with a separate manual, introducing the data reference sheets and electronic resource.

The data sheets were compiled from the comparative data for all the schools for the baseline assessment in 2009 (see Figure 7.7). The comparative data were presented for each of the scales. The sheets supplemented the data from the reports that examined the school results in comparison to the previous year's results. Each school received a data sheet printed back-to-back and laminated.

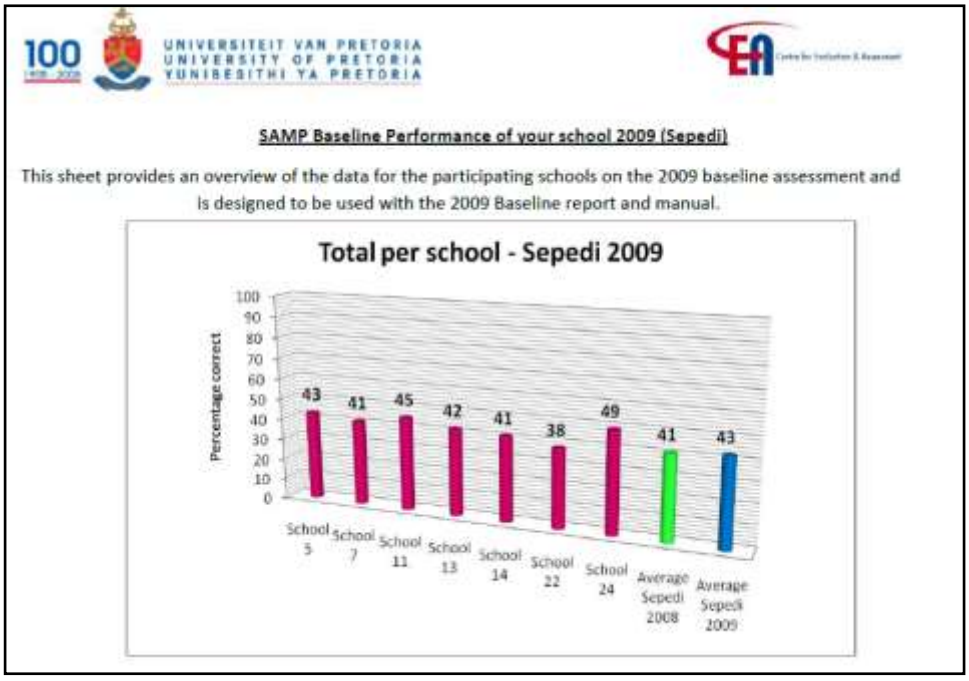


Figure 7.7: Excerpt from datasheet – Prototype III

The electronic resource is discussed in the next section.

7.1.1.4 Electronic Resource

The electronic resource (see Figure 7.8) was produced in the form of a CD or DVD[43] and was provided to each school. Each resource contained a collection of resources for classroom or home use, varying from printable materials to presentations and electronic games. The materials on the electronic resource were a collection of freely available materials from a variety of sources in both South Africa and internationally. Each resource also included an electronic copy of the specific school's report and a manual. Bookmarks to each section of the electronic report allowed readers to click on the section name for easy navigation of the report. The resource also contained a dataset of the school's own data to facilitate further analysis of the data by the school through quick sorting and filtering functions (see Figure 7.9). All resources were evaluated by CEA staff for inclusion and all documents and programmes were loaded on the DVD/CD to ensure schools could use the resource even if internet connectivity was not available.

Figure 7.8: Electronic resource – Prototype III

[43] Schools were contacted to determine if they had access to a computer with a DVD or CD-Rom to determine which format they required

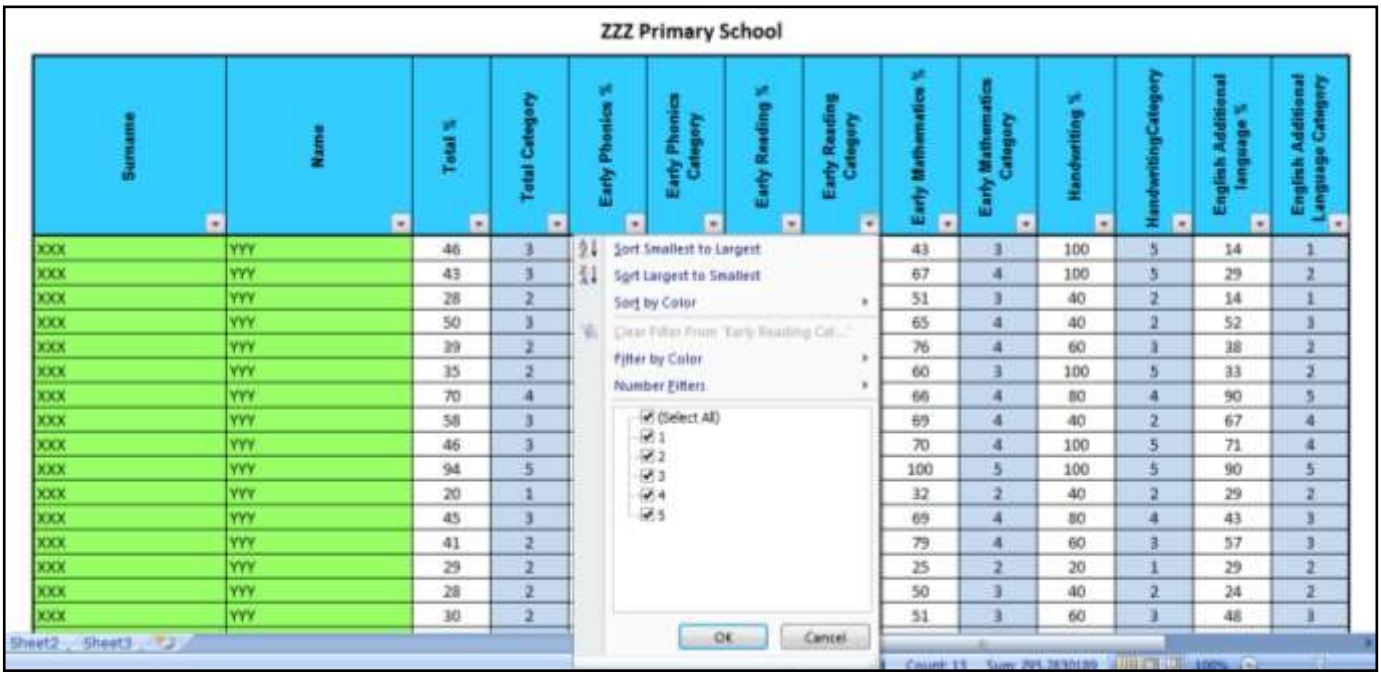


Figure 7.9: Electronic dataset – Prototype III

7.1.2 Formative Evaluation of Prototype III

The formative evaluation of Prototype III took place with two foci. Firstly, the evaluation aimed to examine the practicality and use of the 2009 baseline report that incorporated a separate manual and support CD/DVD with an electronic report and dataset through a questionnaire. The evaluation of the 2009 report was also compared to the 2008 report. Secondly, an attempt was made to observe the transformation of the feedback into planning and action through observations of school meetings, structured reflective journals kept by schools and group interviews with teachers, HoDs and principals to follow up on the observations and journal data.

7.1.2.1 Sampling

All schools in the SAMP sample participated in the ***feedback prototypes***, i.e. all 22 schools in the sample received reports, were invited to the feedback session and received support material and telephonic support if required. The report evaluation questionnaire was handed out at the feedback session, while the school observations, journals and interviews took place at the schools themselves

Sampling for report evaluation questionnaire

All schools who participated in the feedback session were asked to complete the report questionnaire. In most cases each school completed more than one

questionnaire as teachers, HoDs and Principals were in attendance. A non-response analysis was conducted (see Table 7.2), with all but one English and one Afrikaans schools participating in some round of the analysis and three of the seven Sepedi schools doing so. In total, 16 of the 22 schools participated in the report questionnaire, with a total of 28 individual responses.

Table 7.2: Response analysis for Report Questionnaire –Prototype III

School	Grades represented	Language of instruction	Number of questionnaires
School 4	1	English	1
School 5	1, 2	Sepedi	2
School 8	2	Afrikaans	1
School 11	1	Sepedi	3
School 12	R	Afrikaans	1
School 16	1	English	2
School 17	1	English/Afrikaans	1
School 18	1	English	2
School 20	1	English	1
School 21	1	English	1
School 22	1, 2	Sepedi	3
School 23	1, 2	English	4
School 25	1	Afrikaans/English	1
School 26	1	English	1
School 27	1, R	Afrikaans	2
School 28	1	Afrikaans	2
Total			**28**

Sampling for observations, journals and interviews

The sample for this paper consists of three schools that were purposefully selected. The schools were selected based on the criteria of usage and scores improving from previous assessments from each language group. These schools participated in the feedback session and showed good gains from the baseline to the follow-up assessment. One school was chosen per language group (English, Afrikaans and Sepedi).

7.1.2.2 Data Collection

For the report evaluation, all participants in the feedback session were provided with questionnaires to complete. The questionnaires were not only focused on the new 2009 baseline report format, including the separate manual and electronic report, but also had a comparative component in which the schools could indicate if they preferred the 2008 or 2009 report format for each specific component (See audit trail CD/DVD)

The data collection for the three sampled schools employed three different instruments with the data collected in each informing and guiding the development of the following ones. This meant that there was an iterative process between data collection and analysis to gain an in-depth understanding of how the schools use the feedback system. The instruments were:

1. **Non-participant observations (n=3):** schools were observed by the researcher during their usual meetings about the feedback data. Participants ranged from teachers, HoDs and principals and were determined by the usual processes the schools employed.
2. **Classroom journals (n=4):** teachers were requested to fill in a daily semi-structured journal about their use of the feedback in their school. Journals were kept over a four- week period. The guiding questions related to:
    - the influence of the feedback on teacher's planning, practice, communication, individual learner support
    - the accessibility and practicality of the feedback for use
    - barriers to use of the feedback
3. **Semi-structured interviews (n=5):** these were conducted with principals, HoDs and teachers. Principals were interviewed individually and the teachers and HoDs participated in focus groups. The interviews served to follow up on the observations and journals and gain further clarification. The semi-structured questions focused on:
    - the types of data and materials in the feedback system that were used
    - the types of use that took place and the purpose
    - procedures of data-use in the schools (e.g. data paths, different roles of staff, additional training and support)
    - changes in the school attributable to use of the feedback data
    - barriers to use of the data
    - possible improvements in the feedback of the data

The data collected through these instruments were thematically analysed.

7.1.2.3 Data Capturing

Data from the report questionnaire was captured electronically both in textual forms for qualitative analysis and in *Excel* and *SPSS* for descriptive analysis of

frequencies. The qualitative data from the school observations, journals and interviews were captured through field notes from meetings, the journal data and transcription of interview data.

7.1.2.4 Data Analysis

Descriptive statistics were analysed using the frequencies, mean and mode values. Data from the three cases were analysed throughout the data collection process. The analyses informed development of later instruments to clarify information received from earlier data processes. This iterative data collection and analysis process resulted in a rich picture of data-use in each of the schools.

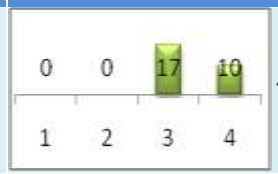
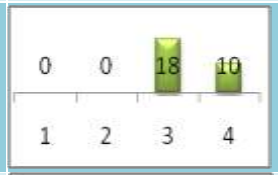
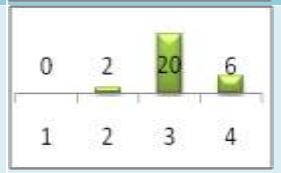
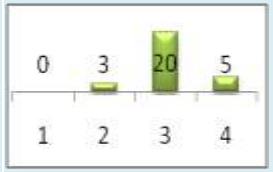
All data were captured electronically and analysed per meaningful unit of text. Codes were generated through inductive process and allocated to each unit of text. Once coding was completed, codes were clustered together in meaningful groups to generate themes. For example, the theme *professional development* includes codes such as *use to increase data-literacy* and *use to identify training needs*. The analysis process was facilitated by the computer aided qualitative data analysis programme *Atlas.ti*.

7.1.2.5 Results and Findings - Report Evaluation Questionnaire

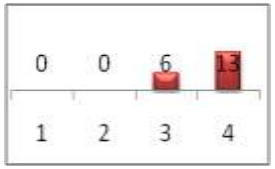
The results of the report questionnaire are summarised in Table 7.3. Overall, rating for the 2008 and 2009 reports were both very high (3.2 and 3.6 out of a possible 4 respectively). The schools however indicated a preference for the 2009 report on all aspects. The following aspects were investigated:

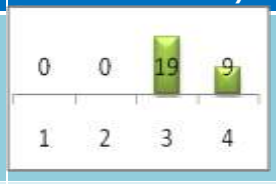
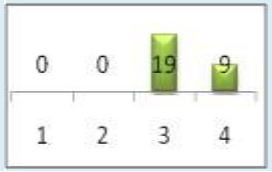
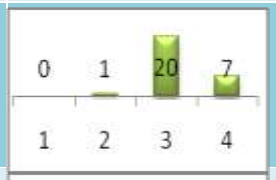
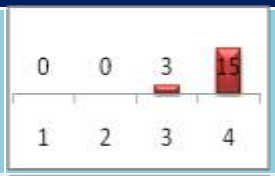
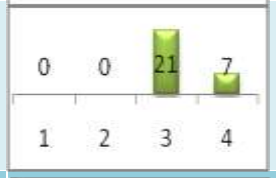
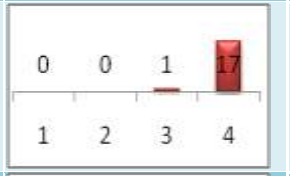
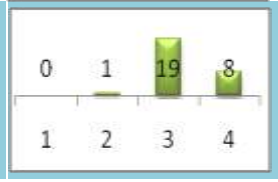
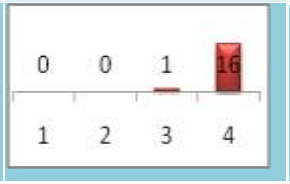
- Usefulness of the information (now incorporated in the manual)
- Presentation of the results
- Support for implementation of results
- Layout and accessibility
- Overall impression
- Overall appropriateness and influence

Table 7.3: Report Evaluation Questionnaire Data

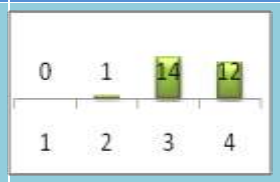
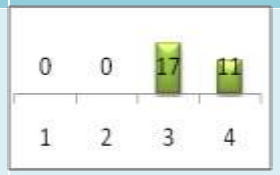
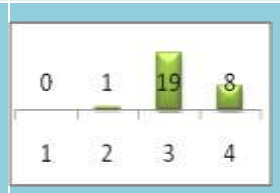
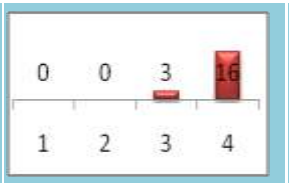
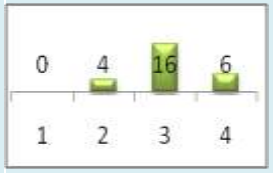
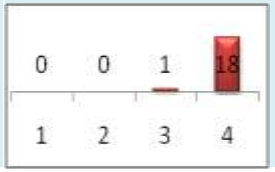
Question	Average rating 2008 1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2008 (1=Poor - 4 = *Excel*lent)	Written comments 2008	Average rating 2009 (1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2009 (1=Poor - 4 = *Excel*lent)	Written comments 2009	Mode preference 2008/2009
Usefulness of the information	**3.3**			**3.5**			
1. Introduction and explanation of project	3.4	 [44]	-	Same for 2008 and 2009			
2. Description of the SAMP assessment	3.4		Very good.				
3. Description of subtests and scales	3.1		-	3.7		-	2009
Presentation of results	**3.2**			**3.6**			
4. Graphs comparing performances of schools	3.1		-	3.6		-	2009

[44] The numbers below the X-axis: 1,2,3,4 represents the different categories on the scale, 1=Poor- 4=Excellent. The numbers above the X-axis represent the frequency of responses for each category. In this example, 13 respondents rated this aspect as 4, or Excellent. Please also note that, not all respondents answered all questions.

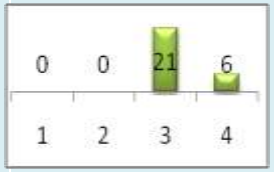
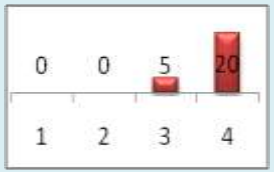
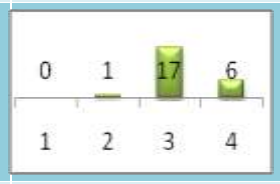
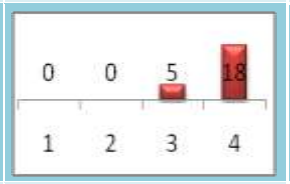
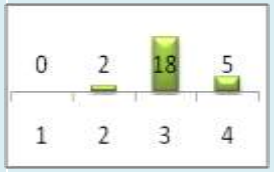
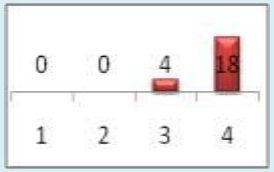
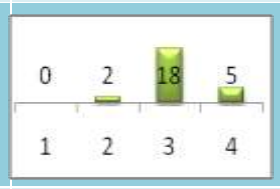
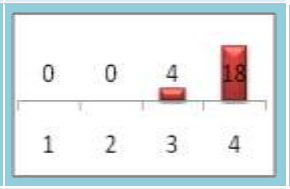
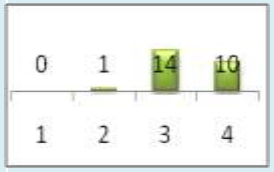
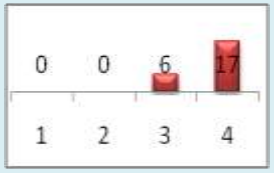


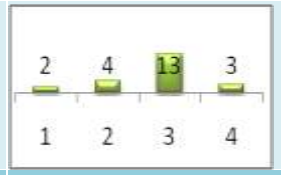
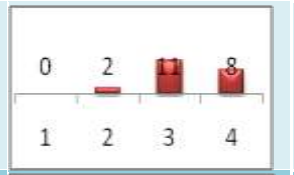
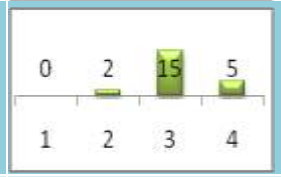
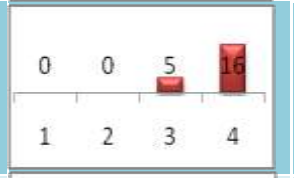
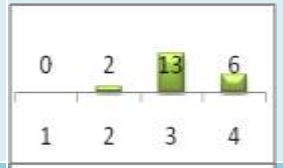
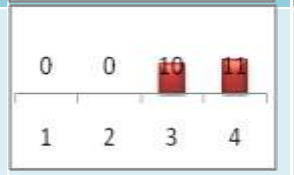
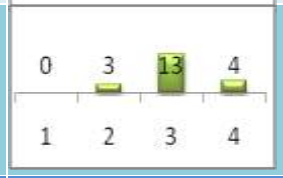
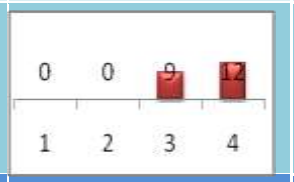
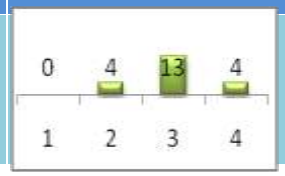
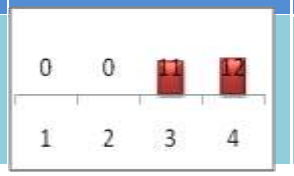
Question	Average rating 2008 1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2008 (1=Poor - 4 = *Excel*lent)	Written comments 2008	Average rating 2009 (1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2009 (1=Poor - 4 = *Excel*lent)	Written comments 2009	Mode preference 2008/2009
5. Graphs showing my school's differential performance	3.3		-	Same for 2008 and 2009			
6. Discussion of graphs and tables	3.3		-				
7. Presentation of individual results in table format (categories)	3.2		*Digitaal sal goed werk indien 'n rekenaar naby is.*[45]	3.8		-	2009
8. Presentation of individual results in table format (%)	3.3		-	3.9		-	2009
9. Distribution of learner performance graphs	3.3		-	3.9		-	2009

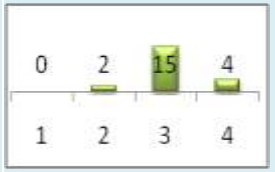
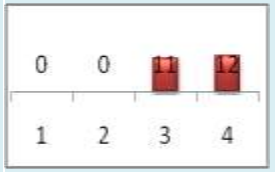
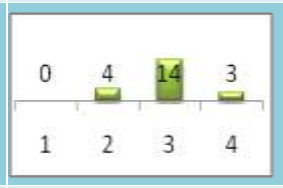
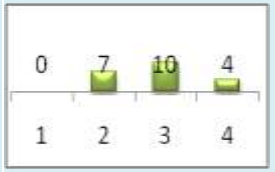
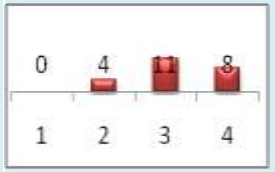
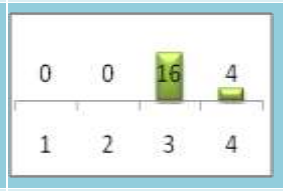
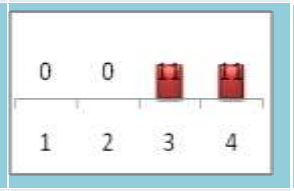
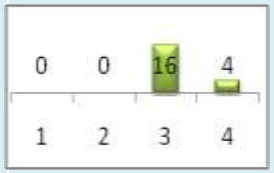
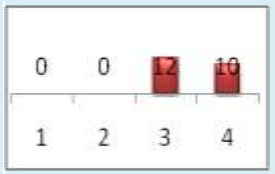
[45] English translation: *"Digital will work well, if a computer is nearby."*

Question	Average rating 2008 1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2008 (1=Poor - 4 = *Excel*lent)	Written comments 2008	Average rating 2009 (1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2009 (1=Poor - 4 = *Excel*lent)	Written comments 2009	Mode preference 2008/2009
Support for implementation of results	**3.3**			**3.6**			
10. Identification of learners at risk	3.4		Very good. Assist in confirming assessment.	**Same for 2008 and 2009**			
11. Identification of exceptional learners	3.4		It works				
12. Conclusion and recommendations	3.3		Thank you	3.8		-	2009
13. Links to and provision of support materials	3.1		-	4.0		*Dankie vir jou moeite!*[46] Thanx!! Very informative and helpful.	2009

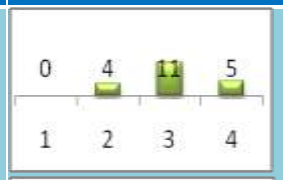
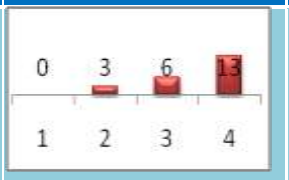
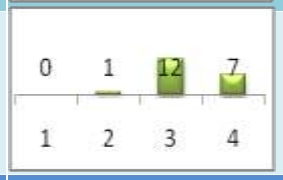
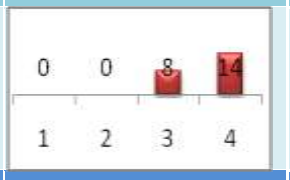
[46] English translation: *"Thank you for taking the trouble."*

Question	Average rating 2008 1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2008 (1=Poor - 4 = *Excel*lent)	Written comments 2008	Average rating 2009 (1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2009 (1=Poor - 4 = *Excel*lent)	Written comments 2009	Mode preference 2008/2009
Layout and accessibility	**3.2**			**3.8**			
14. Ease of finding specific information	3.2		Like it on paper	3.8		Better. Time-saver.	2009
15. Clarity of report	3.2		-	3.8		-	2009
16. Format and presentation of report	3.1		-	3.8		-	2009
17. Support for understanding and interpreting reports	3.1		-	3.8		-	2009
18. Order of presentation of information in report	3.4		-	3.7		-	2009

Question	Average rating 2008 1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2008 (1=Poor - 4 = *Excel*lent)	Written comments 2008	Average rating 2009 (1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2009 (1=Poor - 4 = *Excel*lent)	Written comments 2009	Mode preference 2008/2009
Overall impression	**3.0**			**3.6**			
19. Time between testing and report completion	2.8		-	3.3		Can still improve. Info was not distributed once received at school to grade 1 educators, will advise as info is interesting	2009
20. Overall quality and impression of report	3.1		-	3.8		-	2009
21. Relevance of data in the report to my needs and concerns	3.2		-	3.5		-	2009
22. Amount of data presented in report	3.1		-	3.6		-	2009
Overall appropriateness and influence	**3.1**			**3.5**			
23. Does the report provide required information?	3.0		-	3.5		-	2009

Question	Average rating 2008 1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2008 (1=Poor - 4 = *Excel*lent)	Written comments 2008	Average rating 2009 (1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2009 (1=Poor - 4 = *Excel*lent)	Written comments 2009	Mode preference 2008/2009
24. Does the report help you to assist weak/strong learners in your class?	3.1		-	3.5		-	2009
25. Does the report provide support to address identified problems?	3.0		-	3.5		*Dit bevestig my assessering vir die leerders.*[47] | More!	2009
26. Does the report help with planning and discussions in your school?	2.9		-	3.2		More!	2009
27. Does the report provide you with the opportunity to learn something useful?	3.2		-	3.5			2009
28. Does the report provide your school with the opportunity to learn something useful?	3.2		-	3.5			2009

[47] English translation: *"It confirms my assessment for the learners."*

Question	Average rating 2008 1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2008 (1=Poor - 4 = *Excel*lent)	Written comments 2008	Average rating 2009 (1=Poor - 4 = *Excel*lent	Frequency distribution (count) 2009 (1=Poor - 4 = *Excel*lent)	Written comments 2009	Mode preference 2008/2009
29. Do you trust what is presented in the report?	3.1		-	3.5			2009
30. Do you value what is presented in the report?	3.3		-	3.6		*Dit help om te weet my assessering is op peil.*[48]	2009
Full comments 2009 report	**3.2**			**3.6**			

Question	Theme	Full comments 2009 report
31. What are the strengths of the 2009 report?	**Clarity / Quality**	*Feedback report is mooi uit een gesit en jy weet presies wat aangaan.*[49] | Results are clear for each category tested. | Looks aesthetically good. | Very clear, easy to interpret. | I was impressed with the layout and professional way of report presentation. | I like everything about it. | Thanks for all work from you and team!
	Professional development	*Vind dit leersaam.*[50]
	Support material	More convenient because of the DVD/CD. | *Die CD is waardervol met al die inligting, veral vir die ouers.*[51] | The digital media format of resources and report. Teachers are able to view it electronically. | *Ek hou van die bronne wat beskikbaar is.*[52]| Can't wait to use the CD/DVD!
	Timing	Early feedback.
	Additional variables	Broader parameters. | *Baie uitgebrei.*[53]

[48] English translation: *"It helps me to know my assessment is on par."*
[49] English translation: *"Feedback report is set out nicely and you know exactly what is happening."*
[50] English translation: *"Find it a learning experience."*
[51] English translation: *"The CD is valuable with all the information, especially for the parents."*
[52] English translation: *"I like the resources that are available."*
[53] English translation: *"Very extensive"*

Question	Theme	Full comments 2009 report
32. What aspect of the report would you want changed, added or left out of the report? Please explain.	**Second language assessment**	Test to be done in English as a second language.
	Presentation of data	Preferred the 2008 graphs.
33. In 2009 you received the reports before the feedback session. What difference did this make to you?	**Internal school communication**	Info was not distributed once received at school to grade 1 educators, will advise as info is interesting.
	Earlier opportunity for action	*Jy weet vroegtydig watter leerders benodig meer aandag.*[54] | We could plan and execute our support activities earlier. | Became familiar with it, to discuss with educators in the grade whatever was needed. | Able to reflect on learners results.
	Preparation for feedback session	You come knowing the results. | Better to get it before the meeting. | I was able to be familiar prior to talk and could make opinions and have questions.
	No difference	Didn't really make a difference.
34. What are the weaknesses of the 2009 report?	**Additional variables**	Take learner age into consideration.
	None	I haven't found out any. | Are there any? I think it is great. Thank you for your hard work!
35. What else would you like to have included in the report in future	**Reporting of data**	How do the schools from the different areas compare with each other?[55] Eg. West/North/ East/ Suburbs. A 10 point scale. | I prefer to compare with other schools
	Point of contact	Discuss with the relevant educators not to give it to the principals or clerk.
	More support	How can we help learners who need additional support?
	School identification	I would like to know who the other schools are?
36. Do you have any other comments regarding the reports which have not been addressed above?	**Expansion of Grade coverage**	Are you going to test Gr R in future?
	Happy with project	Thanks for *Excel*lent work and feedback. | It helps us as a guide to verify our judgement of these learners. | I and my staff are truly thankful and we appreciate your hard work.
	Turn-around time	No, Just the report to be earlier.
	Information sessions	An idea would be possibly talk to staff about programme to encourage to take seriously and see value, obviously a fun, short presentation, or a CD to be shown at staff meeting, to encourage enthusiasm.

[54] English translation: *"You know in good time which learners require more attention."*
[55] This suggestion could not be followed-up on as equivalence of the assessment across the different languages has not been established.

Improvement was also noted on every aspect that had changed from the 2008 to 2009 report. Most importantly, the lowest rated aspect of the 2008 report turnaround time from testing to receiving the report was improved from 2.8 to 3.3.

In the open-ended question section responses were grouped according to themes.

The following ***strengths*** of the new 2009 report noted are summarised:

- *Clarity / Quality:* The reports were viewed as clear and of a high quality
- *Professional development:* The reports provided an opportunity for learning and growth.
- *Support material:* The availability of the electronic resource was noted as very positive as it linked the data to support the implementation of changes in the schools. The expanded conclusion section also facilitated this process.
- *Timing:* The faster turnaround time for the 2009 was noted as a strength.
- *Additional variables:* The inclusion of additional variables and data were noted as a strength.

Very few ***suggestions for changes*** were made:

- *Second language assessment:* Some of the non-English medium schools suggested that their learners also be assed in English.
- *Presentation of data:* One school indicated that they preferred the previous year's graphs.

The schools responded to the influence of the ***delivery of the reports prior to the feedback session***.

- *Internal communication in school:* Some schools noted that the earlier delivery allowed for teachers, principals and HoDs to discuss and circulate the information earlier for discussion
- *Earlier opportunity for action:* The data from the report could be acted on even prior to the feedback session
- *Preparation for feedback session:* Some schools indicated that they could be better prepared for the feedback session and have a more purposeful feedback session.

The ***weaknesses*** of the 2009 report were also observed.

- *None:* The schools indicated that they were very happy with the reports and did not experience it as having any weaknesses
- *Additional variables:* One school asked that the ages of learners be incorporated into the data as an additional variable.

In the ***general comments*** section the following ideas were noted:

- *Expansion of Grade coverage:* Some schools recommended the expansion of the assessments to include Grade R.
- *Happy with project:* Most of the schools indicated that they were highly satisfied with the project.
- *Turnaround time:* One school noted that a further improvement of the turnaround time from assessment to report delivery would be welcomed.
- *Information sessions:* It was suggested that information sessions be held at schools to increase awareness of the project and motivate teachers.

7.1.2.6 Exemplary Cases

This section focuses on the data collected through the case studies of the three schools. The observation, journal and interview data for each school is discussed separately for each school. All data were transcribed and thematically analysed using *Atlas.ti*. The three cases are discussed below.

I. SCHOOL A/E

School A/E was situated in a traditionally poor area with predominately Afrikaans learners. Over the previous few years the characteristics of the area had changed and a number of African learners had moved into the area. In response, the school had switched to becoming dual medium with classes in either Afrikaans or English. Socio-economically the learner population was still poor, but little financial support was received from the DoE. As the school was located in an urbanised area with good infrastructure the school was placed in a higher quintile, thus decreasing funding provided by the DoE.

The school principal was committed to a group approach to addressing school issues. He acknowledged the different strengths and weaknesses in staff members and encouraged them to get to know their own personalities and problem-solving preferences during teamwork. The principal felt very strongly that data-based decision-making and planning had to take place and emphasised that all data generated had to be appropriately interpreted and used.

Many of the teachers in the school were still busy with further studies at tertiary institutions. A major focus for the school was ensuring that the new English education component of the school was functioning well. The school often sought alternative and additional funding as many parents could not afford the school fees. The level of parents' education was low and the school struggled to achieve parental participation. Maintaining discipline in the classes was demanding and reduced time on task in the classes.

A/E Data path

At least two Grade 1 teachers from the school attended each feedback session, usually accompanied by the principal or HoD for the foundation phase. Once the reports were received by the school, a meeting was scheduled. Each teacher compared the learner results with their own assessment and list of learners who had been identified for the remedial programme. The HoD was tasked with doing a secondary analysis to determine how learners' pre-schools influenced their performance, in order to

provide feedback to the pre-schools. The principal collated trend data of the school performance based on all previous reports from the SAMP project.

A meeting then took place between the principal, HoD and teachers (see Figure 7.10). The data were discussed and compared to that from other assessments. Tasks were allocated, measurable goals set for improvement and a date set for the next meeting. Follow-up meetings then took place, again in a group format, to monitor progress. The principal also provided the report information to the School Governing Body (SGB) to report on the school's status and progress. The data were also relayed to the School Based Support Team (SBST) to help with planning and support for learners identified as at risk in the feedback. Between the formal meetings teachers and the HoD had informal discussions about progress of individual learners and the success of changes to the curriculum and classroom activities based on the feedback and support materials.

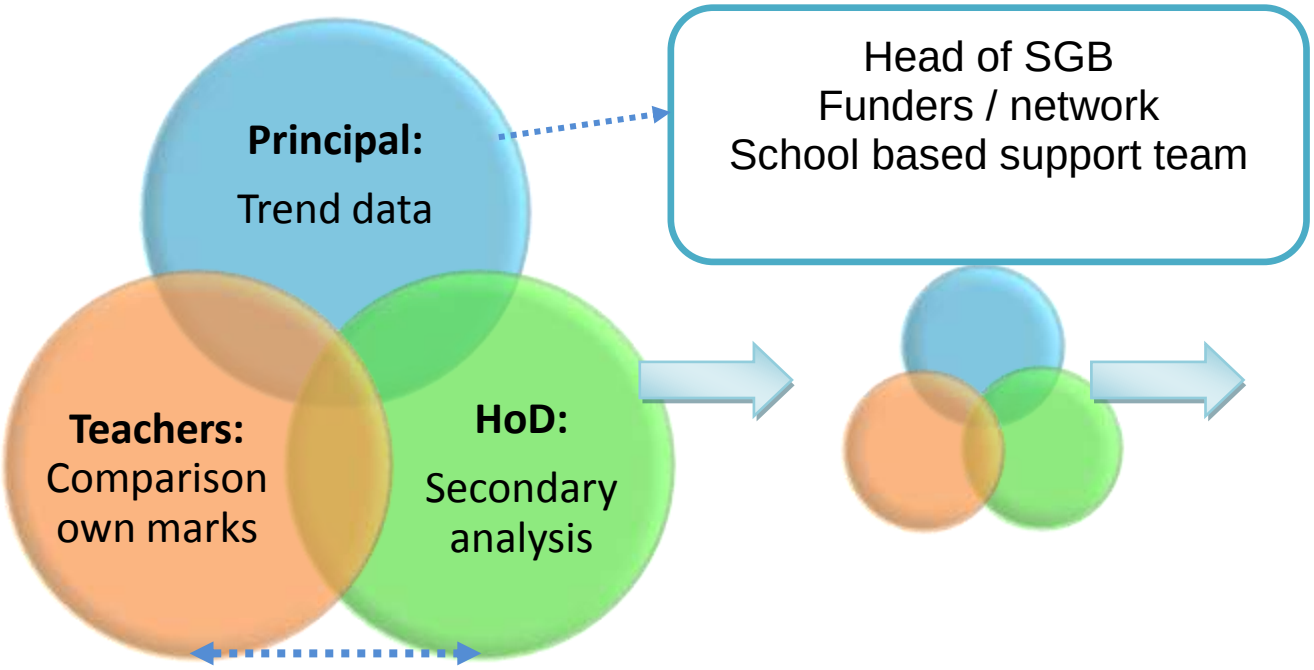


Figure 7.10: School A/E - Team Approach data path

The type of data path observed in School A/E was labelled the ***Team Approach***. The principal (Interview) views this group approach as the most appropriate for his school. *"If I don't drive the process, I don't think the teachers will work with the data to the extent as I wish them to...it is not their natural preference."*

A/E Types of use

In School A/E the data were used to ***support conversations*** between the Grade 1 teachers about individual learners and the school curriculum *"I was concerned about learner X, he did well in your assessments, but struggled in class. So we all [Grade 1 teachers and HoD] sat together and made a plan"* (Interview, Teacher 1). Discussions that are more formal were held with pre-schools in the area to discuss the type of exposure the learners require prior to formal schooling. *"This year we had a meeting with the church pre-school. We talked to them about the skills the Grade 1 learners need for our school and gave them a whole list of tasks they [referring to learners] should be able to fulfil"* (Interview, HoD). Data from the follow-up assessment were also relayed to the Grade 2 teachers to highlight areas in which the Grade 1 learners needed further support: *"… if they [the children] can't do something, the Grade 2 teachers need to know where to start"* (Interview, Teacher 2). The data were used to support requests for further funding and support to the SGB, DoE and potential funders: *"… when I am asking for extra funds for a remedial teacher, occupational therapists, it is something to have in hand"* (Interview, Principal). Where required the data were also used in parent meetings to support teacher evaluations and recommendations. During the observations, the HoD suggested that the teachers use the data to encourage individual learners' parents to become more involved with the learners who were struggling.

The principal and HoD discussed the results from the reports to establish the ***professional development needs*** of the teachers. The data had a direct impact on the ***curriculum development*** for the school's Grade R - Grade 2 classes through the teachers, HoD and SBST (Observations, HoD, Principal and Teachers 1-3)

The principal employed the feedback to meet the ***accountability demands*** to the SGB and the DoE through the mandated IQMS. In School A/E the feedback data were ***triangulated*** with all other assessment data which were available including evaluations from speech therapists and occupational therapists. The data were used to gauge if the assessment standard of the teachers was appropriate and to validate the teachers' identification of

learners who may possibly have failed. *"I used the individual results today to see if I am marking too leniently*" (Journal, Teacher 1). ***Further analysis*** of the data were encouraged in terms of examining trends of the school performance over a number of years and conducting secondary analysis to provide feedback to pre-schools in the intake area of the school. (Observation, agenda and meeting materials)

The data were used for formal ***goal setting*** for individual learners and the school. Measurable goals were set in terms of achieving gains as well as achieving a minimum final score. *"We work for that 10% gain. It motivates us through the year"* (Interview, HoD). The school was concerned with maintaining a high standard relative to other schools with similar characteristics in the sample. The principal was therefore able to ***monitor*** both the gains of learners and performance of the teachers in achieving the goals set in the meetings.

II. SCHOOL E

School E was situated in an area with a pre-dominantly Indian population. Over the previous few years, the characteristics of the school had changed and a number of African learners had started to commute to the school from the township areas. The LOLT in the school was still English, but the majority of the learners were now African and dozens of different home-languages represented in Grade 1.

The school principal was committed to the improvement and further education of his staff, with even the tea-lady involved in tertiary education. He not only encouraged staff to participate in the professional development programme by the DoE, but also arranged for private training opportunities. At the time, the Grade 1 teachers were using a new programme called 'Sheltered Instruction Observation Protocol' (SIOP) to focus on language development throughout all the learning areas to improve learning.

The main challenge facing the school was language development of the learners, most of whom were not first language English learners. As Teacher 3 (Interview) explained: *"You will find all 11 official languages in any class in the*

school and a few extra for a bonus". There was also a high mobility rate of staff as the teachers were often 'head-hunted' by more affluent schools after they have completed their additional training and studies. Parental involvement in the school was limited.

E Data path

The principal took a strong interest in the feedback from SAMP, but rarely attended the feedback sessions himself. Mostly all the Grade 1 teachers, the HoD and sometimes teachers from Grades R, 2 and 3 attended. A formal path for the data in the school was in place referred to here as the ***Cascade Approach*** (see Figure 7.11). The first step was a meeting between the principal and HoD about the feedback. Areas of concern were discussed along with areas of strength and potential improvement. Secondly, the HoD presented the data at a formal meeting with all teachers from Grade R to Grade 3. The principal insisted on all the teachers being involved.

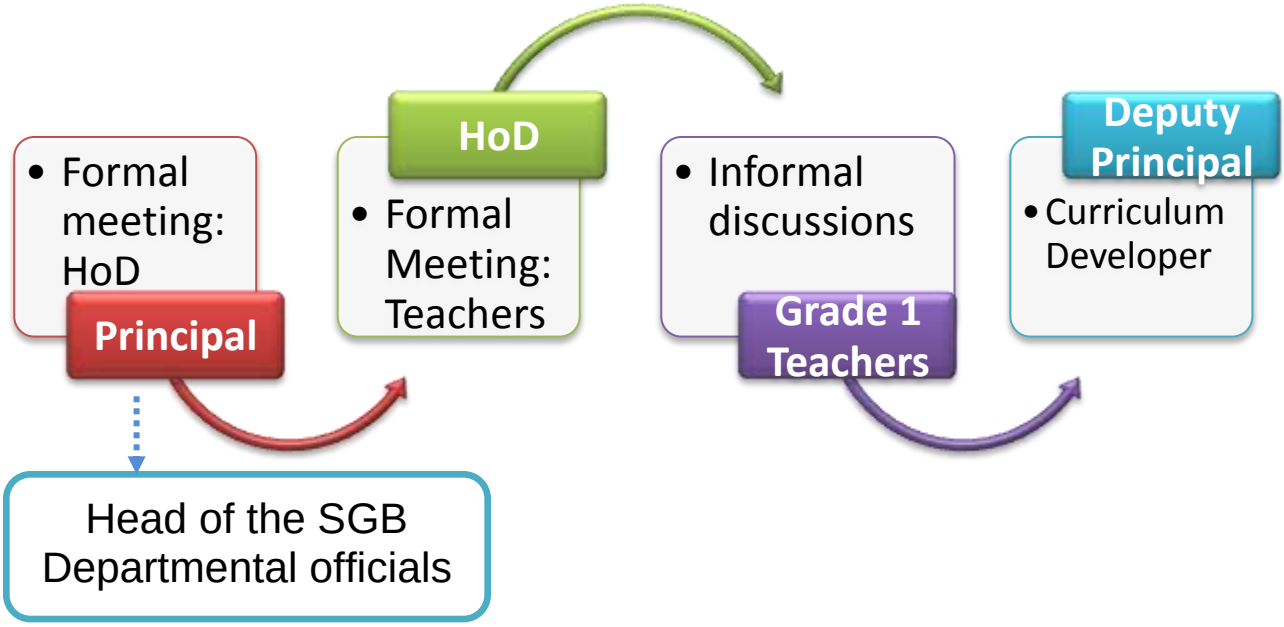


Figure 7.11: School E – Cascade Approach data path

Thirdly, the Grade 1 teachers continued the discussions, mostly informally, about individual learners and changes to the curriculum. The teachers also focused on evaluating whether the new SIOP intervention was proving effective. The curriculum planning for the following year was aligned with the activities recommended in the feedback.

Finally, the data were kept in the deputy principal's office for teachers to access. The Deputy Principal was responsible for curriculum development and integrated the recommendations into the curriculum planning for the following year. The principal (Interview) impressed on his teachers the importance of working with the data, but did not take control of the process: *"I don't get involved with the HoD's meetings with the teachers. It is development for the HoD to interpret the report"*. He did however use the report to report to the SGB and DoE officials to motivate for extra support and funding for the school.

E Types of use

In School E the data were used to ***support conversations*** between the Grade 1 teachers about the efficacy of any new activities, learners in need of support and possible adjustments to the curriculum. *"It is important for us to share what activities are working or not"* (Interview, Teacher 1). The data were used to encourage communication between teachers from Grade R to Grade 3 to facilitate an alignment of goals for the school and coordinate curriculum planning. *"… [A]ll the Foundation Phase and Grade R teachers meet about the feedback, because they will all see these children somewhere down the line."* (Interview, Principal). Data were used to report on the school to the SGB and to garner additional funding from external funders and the DoE. Data were also used upon occasion in parental discussion as an external validation of teacher evaluations and recommendations. "*Sometimes the parents won't believe us that their child is struggling. Then we can show them the marks and say look this was done scientifically by the university, it is not just in our tests that he is struggling"* (Interview, Teacher 2)

The principal and HoD discussed the results from the reports to establish the ***professional development needs*** of the school. Based on discussions, private and DoE training opportunities were identified to address the needs and additional funding was sought if necessary. The principal also saw the process of working with the data as a professional development opportunity for his staff and therefore supported his HoD in interpreting the data during his meeting with her. These data also underwent ***further analysis*** to establish if the SIOP intervention was having an impact on the learners: *"It was good to*

see that the vocab. results increased so considerably, it seems as if the SIOP is working" (Observation, HoD)

These data had a direct impact on the ***curriculum development*** for the whole school through the Curriculum Developer (deputy principal). *"The reports definitely get used… eventually they end up with our deputy principal for curriculum development"* (Interview, Principal). A 15-minute school-wide daily reading period was introduced, based in part on the data from the SAMP feedback. The aim of the reading period was to improve learner literacy across the school. The curriculum development was also impacted directly through the formal teacher meetings of the Grade R- Grade 3 teachers.

The principal employed the SAMP feedback to meet the ***accountability demands*** of the SGB and the DoE through the mandated IQMS. *"See, [shows IQMS file] I even have your documents in my IQMS file" (*Interview, Principal). Data were not viewed in isolation in School E, but ***triangulated*** with classroom assessment data and the principal's classroom observations. The teachers appreciated the opportunity to validate their assessment standards with an external source. This was particularly useful in the difficult task of identifying learners who were at risk of failing half-way through the school year. *"It is difficult to identify the learners, you always wonder… now we can look at the feedback results as back-up"* (Interview, Teacher 3)

The data were also used for formal ***goal*** setting, to evaluate the success of actions based on it. Measurable goals were set in terms of achieving gains as well as achieving a minimum final score. *"The overall score is higher than last year. The score on the Rhyming Words subtest is worrying. The gain is smaller than last year. We said we wanted to increase the gain."* (Observation, HoD). The school was also concerned with maintaining a high standard relative to other schools with similar characteristics in the sample. The principal was therefore able to ***monitor*** both the gains of learners and performance of the teachers in achieving the goals set in the meetings.

III. SCHOOL S

School S was situated in a township area. The LOLT was Sepedi. Learners were predominantly from the immediate area. The socio-economic status of the area was very low and many of the parents were unemployed and generally fairly young.

The school principal was committed to ensuring the school runs smoothly and that the school was an asset to the community. The school was sensitive to the plight of the community and often provide food parcels to hungry families.

The school was involved in dealing with the social problems and poverty the children faced in the home environment. The teachers had a low level of training that the principal was trying to address through professional development activities from the DoE. Parental involvement in the school was poor and learners had little support in the domestic environment.

S Data path

All the Grade 1 teachers attend and participated in the feedback sessions. The principal took responsibility for the use of the feedback and was also usually in attendance. The principal read and interpreted the reports for the teachers: *"I summarise it and tell the teachers what the report says …informally in the hallways or on class visits, also in the meeting after we receive the reports."* This ***Top-Down approach*** is illustrated in Figure 7.12. She also communicated directly to the Grade R and Grade 2 teachers about the feedback as applicable to them. Teachers discussed individual learners and specific suggested activities amongst each other, informally. The Grade 1 teachers mainly examined the report in terms of individual learner results and the conclusion section with the suggested activities.

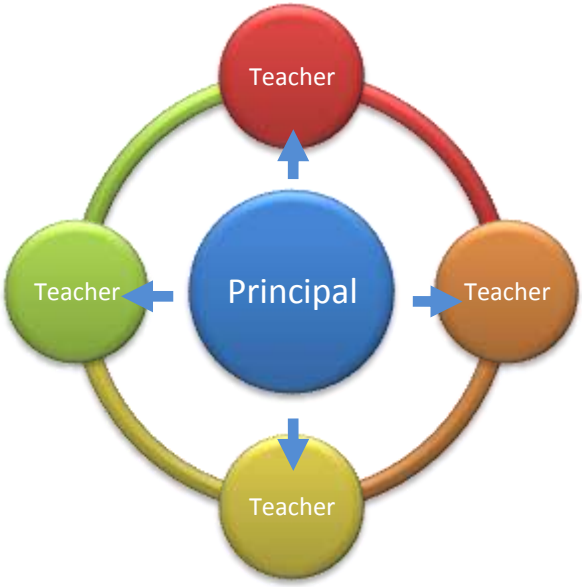


Figure 7.12: School S – Top-down Approach data path

S Types of use

In School S the data were used to ***support conversations*** between the Grade 1 teachers about the various suggested activities, learners in need of support or extra stimulation and possible adjustments to the curriculum. *"We haven't tried these activities [points to suggested phonics activities in report] this may help with the phonics problem, we can try some rap songs"* (Principal, Observation). The data were sometimes used to support conversations with parents about learners experiencing problems or at risk of failure. *"The other day one of the grandfathers wouldn't believe his grandchild was failing, he was giving the teacher a real tough time. Once she showed him the marks from you [referring to SAMP] test, he believed her"* (Principal, Interview).

The principal interpreted the data and explained to the Grade R and Grade 1 teachers what ***curriculum development*** was required. She then also monitored to see if the changes were taking place at classroom level. *"I am in and out of classes all the time, talking to the teachers and seeing if they are doing what we talked about"* (Principal, Interview).

The principal supported the teachers in using the data to ***triangulate*** with their own classroom assessment standards. In this way the teachers could establish if they are marking at an appropriate level. The data were also

compared to the learners who were identified to the DoE as being at risk of failure.

The principal facilitated formal ***goal-setting***, operationalised in terms of gains, relative achievement to other schools and minimum final results in the feedback reports. *"As long as we are showing a gain, I am happy"* (Principal, Observation). These goals made it possible for the principal to ***monitor*** learner gains and teacher performance. The goals also served as a motivational factor for the teachers. *"Just look how happy they are when they hear the results"* (Principal, Interview).

7.1.2.7 Discussion and Design Guidelines

The focus of this chapter was predominantly on how data from the feedback system are used in the schools. Brinko (1993) suggests that when designing a feedback system the following be borne in mind:

- Gathering from multiple sources, by oneself as well as by others.
- Information should be seen as accurate, credible, well-intentioned, descriptive and specific and so more likely to be effective.
- A variety of modes should be used to feed back information.
- Feedback should be seen as a process and not a 'quick-fix' then feedback would be more effective.
- Participants should be part of the process and select the way in which feedback is conveyed.
- Negative feedback is more effective when it is sandwiched between positive information and is self-referenced rather than norm-referenced.
- An opportunity should be provided for response and interaction.

In this research these elements have been an important point of departure and have been successfully implemented. Clearly, the concept of use is difficult to untangle as there seems to be evidence of overlapping of different types of use depending on the purpose. The idea of adoption implementation in which the dissemination of data, evaluation of applicability and incorporating the data

into practice (Brown & Rodger, 1999) is used, is clearly illustrated by the three case studies. While the three schools used three very different approaches to data-use , all however used the data for multiple purposes, as illustrated in Table 7.4 (For a full discussion on use of feedback please refer back to Section 3.4).

Three distinct approaches to data-use that appear to be effective were identified and explored in this cycle: Team, Cascade and Top-down. The most appropriate and effective approach of use may depend on the culture of the school, school leadership approach, level of teacher development, context and current level of functioning of the school. A more advanced, sophisticated approach to data-use may not always lead to improvements, but may be disheartening and inappropriate in a certain context.

There are, however, certain commonalities in the approached to effective data-use illustrated above. In all cases the principals valued and emphasised the importance of data-based practise and planning. Formal time and space was created to work with the data. Multiple role players were involved in the data analysis, interpretation and application process, from teachers to HoDs and principals. Target-setting was used to motivate teachers and monitor progress in the schools. School data were never viewed in isolation, but interpreted in light of other sources through triangulation. In all cases the data were used to facilitate conversations about the curriculum and individualised support with role players such as other teachers from different grades, parents, the SGB and the DoE. Feedback was also provided with links to support material to provide a stepping-stone to action. From the case studies it appears that an effective feedback system should thus try to establish or encourage these conditions for data-use. Data must also be provided in such a manner that it can meet the needs of different schools at various levels of data-use sophistication.

Table 7.4: Summary of schools use of the SAMP data

USES		School A/E  Team	School E  Cascade	School S  Top-down
Supporting conversations	Teachers:			
	Grade R/Pre-school	×	×	×
	Grade 1	×	×	×
	Grade 2	×	×	×
	Grade 3		×	
	School Governing Body	×	×	
	Departmental officials		×	
	Professional development providers	×	×	
	Funders	×	×	×
	Parents	×	×	×
Professional development	Increasing data-literacy	×	×	
	Identifying training needs	×	×	
Curriculum development and planning	Grade R	×	×	×
	Grade 1	×	×	×
	Grade 2	×	×	
	Whole School	×		
	Curriculum developer		×	
	School Based support team	×		
Meeting accountability demands	IQMS	×	×	
	SGB	×	×	
Triangulation	Own assessment standards	×	×	×
	Possible failure identification	×	×	×
	Other data	×		
Further analysis	Evaluation of interventions		×	
	Secondary analysis	×		
	Trend data	×		
Goal setting	Gains	×	×	×
	Final results	×	×	×
	Relative	×	×	×
Monitoring	Teacher performance	×	×	×
	Learner gains	×	×	×

Design guidelines from Report Evaluation Questionnaire

The design guidelines for the report component of the feedback system were collected through an evaluation questionnaire that compared the functioning of the 2008 and 2009 reports. The following design guidelines were generated from the data.

1. ***Improved turnaround time is important*** in increasing the relevance, usefulness and efficacy of the feedback. This may be accomplished by employing ***ICT tools*** to decrease turnaround time through automation of certain processes.
2. ***Feedback must be linked with resources and suggestions for action***, the presence of links to support material facilitate the use of feedback. Resources, however, do not ***need not be developed anew***, as collecting and providing access to available resources in an organised, centralised manner is sufficient.
3. ***Schools are interested in secondary and additional analysis*** as indicated in their request for the addition of variables in the analyses. Schools can be provided with tools to conduct their own additional analyses and therefore additional analyses need not be conducted externally. If ***electronic datasets*** are provided in a commonly used package, ***schools can conduct additional analyses independently***.

Design guidelines from Case Studies

The observations, journals and interviews conducted for the three case studies provided a deeper understanding of how schools use data. The data were employed to construct design guidelines on how to facilitate use of the feedback data in schools:

1. The feedback system should have an ***underlying rationale*** and design that solidifies the link between understanding the feedback and emphasising the importance of ***data-based practise and planning. Links to resources*** and support materials are essential to facilitate action and emphasise this rationale.

2. The system should be designed in such a way that the ***minimum time is spent on understanding*** the data, to allow users to spend time on planning and taking action based on the feedback.
3. The feedback system should allow for ***target-setting*** and a way to ***monitor*** progress in attaining the goals.
4. The system should ***encourage triangulation*** of data, by acknowledging the limitations of a single method of data collection and acknowledging the importance of other modes.
5. Feedback must be easily ***linked to the curriculum*** to facilitate conversations, planning and action.
6. Data must be ***differential and detailed*** enough to be practical for a variety of purposes, as schools need to use the data in a variety of contexts.
7. Data must be ***presented in different manners*** and levels of sophistication to be of use to schools that may be at different levels of development and data-literacy. ***Providing a dataset*** that can easily be used for secondary and additional analysis is essential to fulfil the needs of schools which are more data-literate.

7.2 Conclusion

In this chapter the last cycle of the Prototyping Phase of the design research process was documented. The emphasis of the cycle was on transforming the conditions of use into action in schools. This was achieved by examining the current use and processes employed by schools in planning and taking action based on the feedback system. Three different approaches to data paths were identified: Team, Cascade and Top-down. The design guidelines from this cycle informed the development of the fourth and final prototype, Prototype IV to be documented in the next chapter. The next chapter therefore focuses on the final phase in this thesis, namely the Assessment Phase that examined the overall functioning of all the components as a complete system.

CHAPTER EIGHT
Assessment Phase: Cycle 4

> [A feedback Intervention that] provided for a familiar task containing cues that support learning, attract attention to feedback-standard discrepancies at the task level and is void of cues… that direct attention to the self …is likely to yield impressive gains. (Kluger & DeNisi, 1996, p. 278)

This chapter takes a closer look at the research design and results for the final design cycle. This last cycle is described as the Assessment Phase (Plomp, 2009) of the design research process. This phase examines the functioning of the elements of the feedback system, as well as the global functioning of the integrated learner performance monitoring feedback system. The Assessment Phase is discussed with reference to the specific research questions and evaluative foci, Prototype IV, the research design, findings and design principles from the phase.

8.1 Research Cycles

Figure 8.1 demonstrates the design and evaluation interactions for this research in detail. The pre-existing system and feedback prototypes are illustrated in blue and the evaluation activities are illustrated in green. As stated previously, every full design cycle consists of the prototype adaptation followed by implementation and corresponding formative evaluation of that prototype. The Assessment Phase consists of the final design cycles in this research. The previous cycles are discussed in Chapters 5-7.

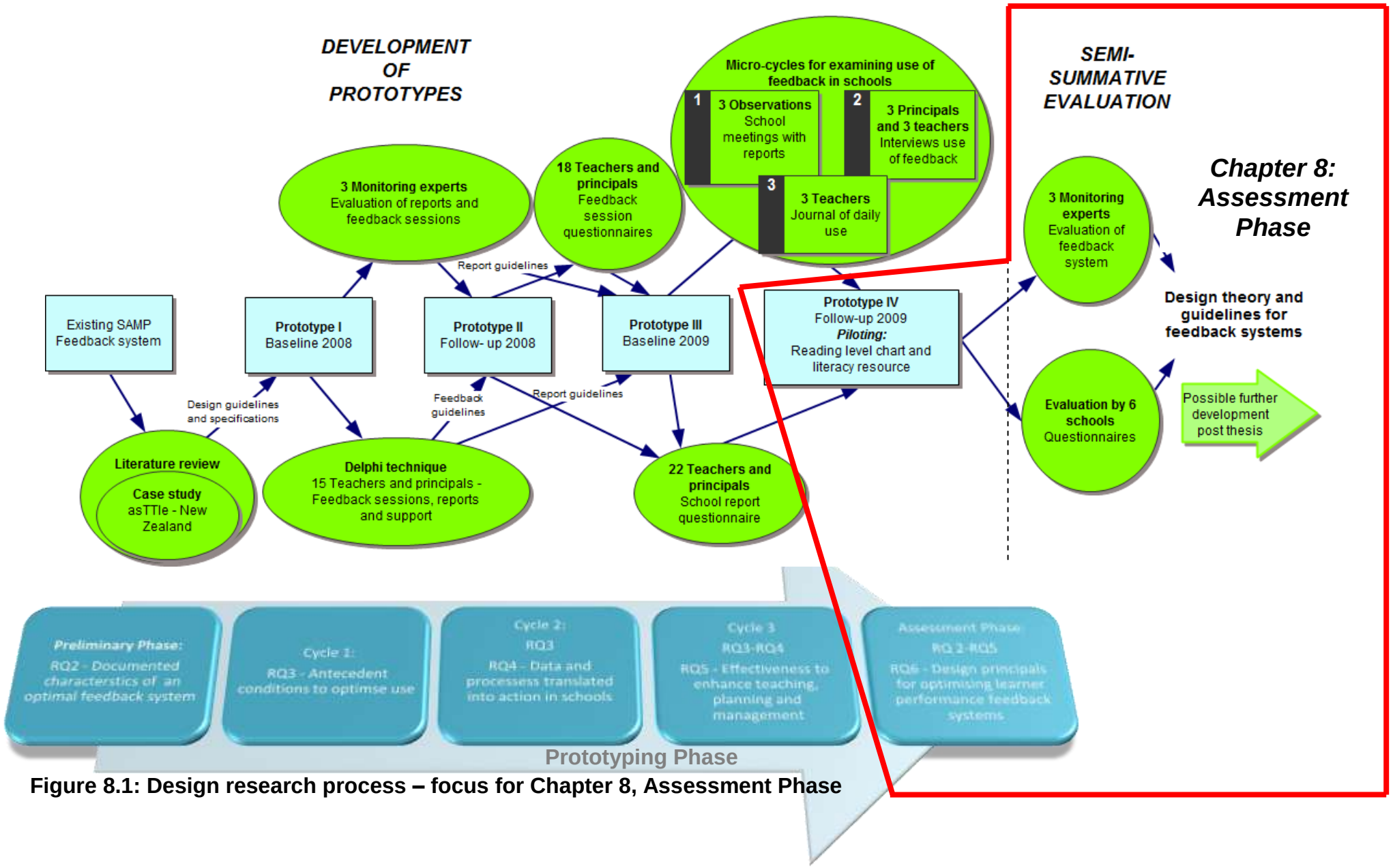


Figure 8.1: Design research process – focus for Chapter 8, Assessment Phase

8.2 Cycle 4 (Prototype IV – Follow-up 2009)

The final cycle incorporated the learning from all the previous cycles in order to evaluate the functioning of the feedback system as a whole. It therefore consists of a semi-summative evaluation (Plomp, 2009) of Prototype IV. The evaluation was conducted by means of reports from expert evaluators, a questionnaire for teachers and a questionnaire for school management.

This cycle focused on ***research questions*** 3-5:

3. *What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?*
4. *How do schools use feedback?*
5. *How effective is the feedback system in enhancing classroom practices, management and planning activities?* ***(focussing on expected and expressed efficacy)***

The evaluation information served to provide ***design guidelines*** relating to the development of a functioning feedback system that facilitates use of the feedback.

4. Completed intervention: This stage was aimed at assessing the functioning of the system as a whole and was achieved through the semi-summative evaluation of Prototype IV. In this evaluation, all components were evaluated along with the interaction of the components and their functioning as a whole.

The evaluation in this cycle focused specifically on the ***evaluative foci*** of practicality and efficacy:

3(a) *Expected practicality:* The system as a whole is expected to be usable in the settings for which it has been designed and developed, i.e. schools in the Tshwane area. Therefore, the focus is on whether the overall feedback system is understandable and helpful for the schools in informing action.

3(b) *Actual practicality:* The actual practicality, relates to how accessible, understandable and usable the system is for the school environment it was designed. This was examined through the participants reports of the practicality of the feedback system in their setting.

4(a) *Expected efficacy:* Using the system is expected to result in the desired outcome of helping users to take action and plan for improvement on the individual, classroom and school level. The expected efficacy focuses on whether the users perceive and expect the feedback system to achieve this outcome.

4(b) *Actual efficacy:* The actual efficacy of the feedback system was also explored but only to a limited extent. Teachers and management respondents provided information on how effective they found the system for planning and action in schools. Actual efficacy of the feedback system was examined through observations, interviews and journals in Chapter 7 and the users' views of the actual efficacy are examined in this chapter through the questionnaires. In order to establish actual efficacy fully, causal relationships between the feedback system and changes in the learner results would have to be established. This exploration was not part of the design.

In the following section, Prototype IV of the feedback system, as it was developed and implemented in the final cycle is introduced.

8.2.1 Prototype IV – Follow-up 2009

The sample of 22 schools (Afrikaans, English and Sepedi) was maintained for the follow-up 2009 assessment. Therefore, 1,569 Grade 1 learners from the baseline assessment were assessed in the follow-up for 2009. There was a drop out rate of 7.7%, due largely to learner absenteeism and migration. All these schools participated in the feedback system.

All the schools in the sample received paper-based reports, that were delivered prior to the feedback session. The report and instrument manual formed separate resources, which maintained the format used in the baseline. Teachers, principals and HoDs were invited to attend the feedback session. Schools were also invited to contact the CEA if there were any questions or any help and support were needed with interpretation and planning. At the feedback session each school received a DVD or CD (electronic resource) containing freely available literacy, phonics, numeracy and mathematics materials. The DVD/CD also contained an electronic copy of the specific school's report as well as an *Excel* dataset for the school to use

for further analysis. The electronic resource was further supplemented with the launch of a support website, which contained the same, as well as additional resources to those on the DVD. The website could be accessed by schools, parents and learners alike. The components of Prototype IV are shown in Table 8.1.

Table 8.1: Prototype IV – Follow-up 2009 components

Component	Description
Paper-based follow-up report	Produced for each school individually and delivered to each school a week before the feedback session
Paper-based instrument manual	Delivered to schools with report
Follow-up feedback session	Principals and teachers from all participating schools invited to the University of Pretoria
Electronic resource	Electronic collection of literacy, phonics, numeracy and mathematics materials produced on CD/DVD and provided to each school at the feedback session. Included an electronic copy of the report (PDF) with a learner dataset for the school in *Excel*
Website	Links and electronic resources for literacy, phonics, numeracy and mathematics. More extensive than electronic resources with web-access for schools, parents and learners.
Telephonic, written and face-to-face communication	On an ad hoc basis as required

The changes to the various components of the feedback system are discussed in the following sections.

8.2.1.1 Reports

The follow-up report reflected the structural changes of the baseline report. The adapted structure of the reports was as follows:

1. Introduction to the SAMP project
2. Results per school ***(updated)***
3. Individual results (category scores)
    a. Notes on interpretation
    b. Learner results table
    c. Learners at risk
    d. Exceptional learners
4. Conclusion and recommendations
5. Individual results (percentage scores)

The content and structure of the reports remained the same from the baseline to the follow-up with a particular focus on facilitating interpretation and use. The results per school were aimed mainly at a management and curriculum level for schools to examine their overall performance. The comparison to other schools also allowed schools to examine their relative progress and determine specific areas which were problematic across the grade for curriculum planning. The individual result section helped identify learner strengths and weaknesses, thus facilitating differential teaching practices. The conclusions summarised and interpreted the overall data for each school with concrete examples of activities that would be appropriate to address the areas of concern.

The turnaround time for report delivery was improved by expanding and refining the automation process from the baseline assessment. The automation process also influenced the representation style of the graphs. The number of bars on the graphs was reduced to three, with only the school averages for 2009 and 2008 as well as the language group average for 2008 being reproduced (see Figure 8.2). This is different from the 2008 reports in which the results for each school were presented by bars in these graphs. This new representation style had the added benefit of making it easier for schools to compare their scores over a two-year period to examine any trends.

By presenting comparative data, the reports contributed to the enlightenment of schools about the functioning of schools in the Tshwane area. The different presentations of the data provided an opportunity for the users to improve their data-literacy. The specific conclusions and recommendations provided modelling of how schools can interpret data and use it in schools, thus providing an opportunity for process use of the data.

A stack bar graph was used to show gains and losses on each of the bars for items that remained consistent from the baseline to the follow-up assessment (see Figure 8.2). The gains score provided a clearer picture of the value added by schools than outcome scores alone. The baseline score for 2009 was therefore represented in green and the gains or losses added in pink. The presentation style of the graphs was also updated with the *2007 Office* package to have a more professional appearance. The vertical axis was also pre-set to a scale of 0-100 to ensure consistency throughout the report.

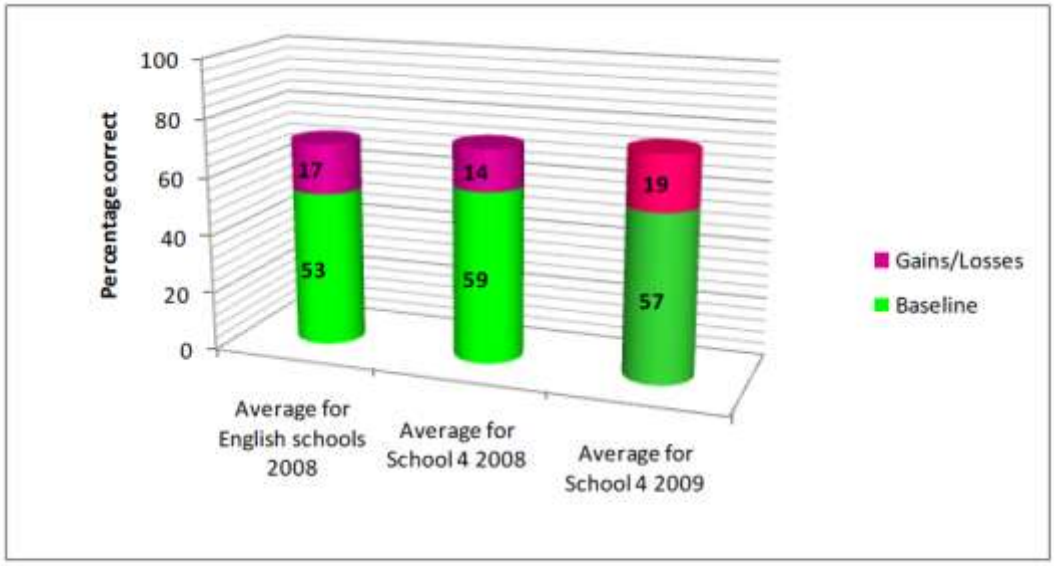


Figure 2: Difference in performance on all items that remained consistent from the Baseline to the Follow-up assessment compared with 2008 data – English Medium of instruction[2]

Figure 8.2: Difference on all items that remained the same from Baseline to Follow-up – Prototype IV

The report was also saved as a PDF document. The electronic version of the report was saved on the electronic resource for each school to allow printing of extra full colour reports if required. Schools could also make use of the report in electronic format as bookmarks were inserted for point-and-click navigation.

8.2.1.2 Instrument Manuals

All the schools in the sample received a separate instrument manual delivered with the reports. The instrument manual format from the baseline assessment was maintained, with added explanations for interpretation of the new follow-up graphs. The structure for the instrument manual was:

1. Introduction
2. Description of baseline subtests
3. Description of follow-up subtest
4. Description of the English Additional Language Assessment
5. Quality of Data
    a. Validity
    b. Reliability
6. Notes on Interpretation ***(expanded)***
    a. Assessment results for the school overall

   b. Individual learner results
   c. Conclusion
7. Summary
8. References

The instrument manuals not only provided tools to interpret the data in the reports, but through the examples and explanations increased the user's data-literacy through conceptual learning. As the instrument manuals provided support for interpreting and using the data, they helped make it more feasible for schools to act on the data, thus influencing the process of feasibility testing as discussed in the conceptual framework.

The format of the follow-up feedback session was similar to that of the baseline. The feedback session is discussed below.

8.2.1.3 Feedback Sessions

All schools were invited to attend the feedback session and had received the reports prior to the session. At the session each school received an electronic resource and data reference sheet for the 2009 follow-up. This showed the school comparative data for 2009 that were not covered in the reports (see Figure 7.7). The feedback sessions took one hour and consisted of a presentation by the project leader, followed by an introduction to the website. This was followed by a certificate presentation to the participating schools by one of our Dutch partners. This feedback session was very well attended with 20 of the 22 schools represented and more than 43 staff members from the schools in attendance. The feedback session pertained to:

- A short introduction to the project
- Developments in 2009 ***(new)***
- A overall description of the assessment
- Presenting the new report format
- An example of how to make sense of the data in the new format
    - Linking interpretation to action and planning ***(new)***
- Introducing the website ***(new)***
- An opportunity for discussion and questions

- Certificate ceremony ***(new)***

Handouts of the presentation were provided along with the 2009 follow-up data sheets and refreshments were served. The data sheets were compiled from the comparative data for all the schools for the baseline assessment in 2009 (see Figure 8.3). The comparative data were presented for each of the scales. The sheets supplemented the data from the reports that examined the school results in comparison with the previous year's results. Each school received a data sheet printed back-to-back and laminated.

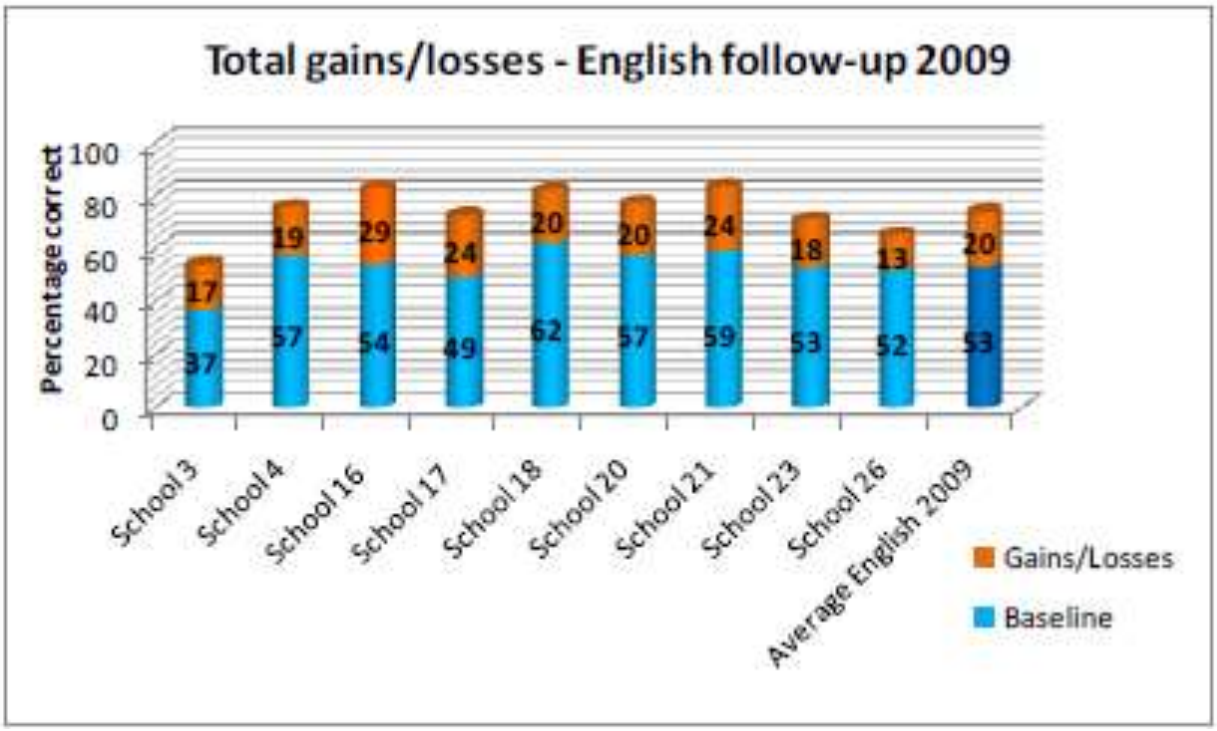


Figure 8.3: Datasheet Excerpt – Prototype IV

The structure of and approach to the feedback sessions provided a valuable opportunity to strengthen the relationship flux characteristics. The two-way communication and reporting of progress as well as discussion of future plans showed a respect of the schools' data users, acknowledging the role of the users as full participants and encouraging a sense of ownership. Where the feedback on the planning and progress reflected user input has been incorporated, sense of ownership was further strengthened. Careful attention to logistical matters and the manner of addressing the school users provided an opportunity for showing respect for the participants and building trust. Support and discussions on how to interpret and use the data helped to improve the users' data-literacy and therefore their ability

to use the data instrumentally. At the same time, the feedback session provided an opportunity for schools to network with each other and form a community of users who can provide support and share experiences. These aspects could however only be accomplished because of the non-judgemental, constructive atmosphere established at this and previous sessions. Feedback sessions provided support and feedback in a manner that did not threaten any school and invited collaboration and participation.

8.2.1.4 Website

The support website (www.samp.schools@officelive.com) was developed to supplement the electronic resource DVD/CDs for schools. The website (see Figure 8.4) has additional resources updated regularly by the CEA staff. The website also has links to other resource sites, downloadable games, classroom and printable resources. It has sections introducing the project and the CEA as well as quick links to contact the CEA staff or to access support for the website.

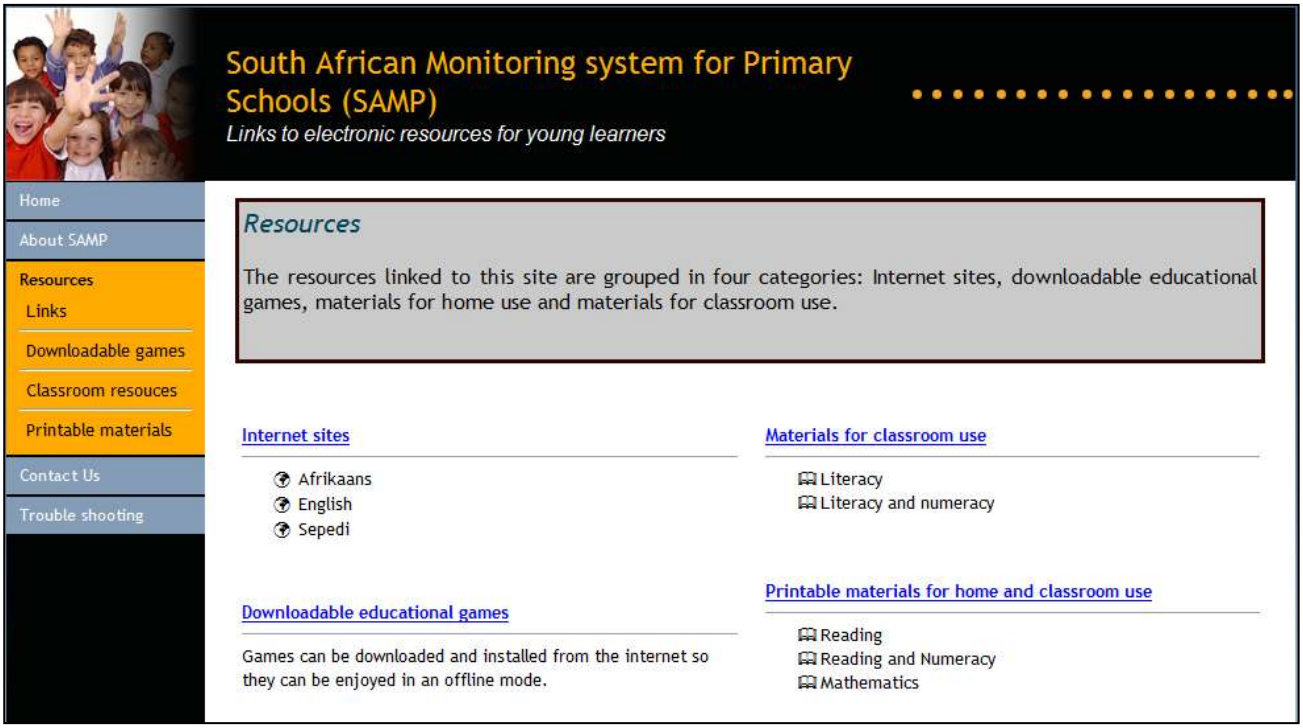


Figure 8.4: Screenshot of the SAMP resource website

The website was created for learners, guardians and schools. Therefore, schools could refer parents and learners to the website to access resources at home to help support the learning taking place in the school environment. In 2009/2010 the website received 149 unique visitors viewing a collective 626 pages.

The resources from the site made it easier for schools to use the feedback appropriately to address identified issues, thus making it more likely that feedback would pass the feasibility test in order to be used instrumentally. The 24-hour accessibility of the site also allowed the schools to function more autonomously with the feedback and increased their sense of ownership of the process. The site also provided an opportunity for the users to gain enlightenment about the use of ICT resources for classroom practice and planning.

8.2.2 Semi-summative Evaluation of Prototype IV

The semi-summative evaluation of Prototype IV was aimed at examining the functioning of the overall feedback system as a whole. This evaluation was conducted by both expert evaluators and school users. The expert evaluators (Dutch partners introduced in Section 6.1.2.1) were asked to evaluate the overall feedback system and write short evaluation reports. While two separate, comprehensive questionnaires were supplied to each school, one for the teachers and one for the management (principals, HoDs and head teachers). The focus of the questionnaires were to determine how the feedback system as a whole was functioning and is referred to as a semi-summative evaluation as further development of the feedback system may still take place (Plomp, 2009).

8.2.2.1 Sampling

The evaluation reports were completed by the two Dutch experts involved in the research project. Each school in the SAMP project was provided with two questionnaires, one to be completed by a teacher and one to be completed by someone from a management perspective e.g. a principal, HoD, or head teacher. A response analysis was conducted (see Table 8.2) seven English schools, four Afrikaans schools and five Sepedi schools completed the questionnaires. In total 16 of the 22 schools participated in the semi-summative evaluation. A total of 27 individual questionnaires were received: 14 teacher questionnaires and 13 management questionnaires.

Table 8.2: Response rates for the Teacher Questionnaire and the Management Questionnaire

School	Language	Teacher Questionnaire	Management Questionnaire Respondents	
			HoD/Head Gr1 Teacher	Principal
3	English	1	1	-
4	English	-	1	-
6	Afrikaans	-	1	-
7	Sepedi	1	1	-
8	Afrikaans	1	1	-
11	Sepedi	1	-	-
12	Afrikaans	1	-	-
14	Sepedi	2	-	1
16	English	1	-	-
18	English	1	1	1
20	English	1	1	-
21	English	1	1	-
22	Sepedi	1	-	1
23	English	1	1	-
27	Afrikaans	1	-	-
24	Sepedi	-		1
16		**14**	**9**	**4**

8.2.2.2 Data Collection

The two expert evaluators were approached during their visit to the CEA to write an evaluation report. Both evaluators were provided with a full set of reports, manuals, electronic reports, website address, agendas, data sheets and supporting documentation. One of the evaluators also attended and participated in the follow-up feedback session in 2009. The evaluators were requested to evaluate all the different components of the feedback system and provide suggestions for improvement. The evaluators produced two short reports on the feedback system, focussing mainly on suggestions for improvement. This was followed up with the evaluators telephonically and by e-mail. The evaluators indicated that they thought the system was functioning well and that they therefore focused their reports on any further suggestions for improvement.

The two final evaluation questionnaires for teachers and management were designed to address research sub-questions 3-5. Both these questionnaires therefore had sections relating to:

- *Sub-question 3: What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?*
- *Sub-question 4: How do schools use feedback?*
- *Sub-question 5: How effective is the feedback system in enhancing classroom practices, management and planning activities?*

The teacher questionnaire focused on a classroom perspective, while the management questionnaire held an overall management perspective. The questionnaire consisted of Likert scales with space for comments and open-ended questions (see audit trail DVD).

The two questionnaires were faxed, e-mailed or hand delivered to schools depending on the availability and functioning of the schools' communication infrastructures. Telephonic or face-to-face support was provided by CEA staff (four schools) if there were any difficulties or if the faxes to or from the schools were unclear. Most schools faxed the questionnaires back to the CEA, but four of the questionnaires were collected as the schools had difficulty with their fax lines. Telephonic follow-ups were conducted where responses were unclear. The questionnaires were sent to schools after completion of the 2009 SAMP cycle and after schools had received all materials and attended the feedback session.

8.2.2.3 Data Capturing

The reports from the two evaluation experts were received electronically and stored as such for analysis. Data from the report questionnaires were captured electronically both in textual forms for qualitative analysis and in *Excel* and *SPSS* for descriptive analysis of frequencies. All data sets and outputs are presented on the audit trail DVD.

8.2.2.4 Data Analysis

The data from the expert evaluators' reports were collated and summarised. Descriptive statistics were examined, including the frequencies and mode values. The analysis, including histograms providing an illustration of the frequency distribution of responses are presented in Table 8.3 - Table 8.14. Mean values were also presented for illustrative purposes, although this is not usually the practise with

non-parametric data. The qualitative responses were also grouped according to themes and some selected quotes are represented in the tables for each question or comment on an item.

8.2.2.5 Results and Findings - Expert Evaluators' Reports

The overall evaluation of the feedback system by the Dutch evaluators was positive.

> From the reactions of the representatives of schools, whom I met during this visit, I am convinced that the linked monitoring and feedback system is very relevant for schools and teachers. In other words, the project has laid the fundament of a unique contribution to South African education by developing a good feedback system linked to the school-based monitoring system …SAMP. (Evaluator 1, 2 April 2010)

> The manuals, baseline feedback reports and follow-up feedback reports are generally well structured and well written. They also contain sensible and easily accessible information as this is presented both in writing, in graphics and in numbers. (Evaluator 2, 17 February 2010)

Some comments and recommendations were made about the different aspects of the feedback system in the report by Evaluator 1 (2 April 2010) and Evaluator 2 (17 February 2010). The comments and recommendations related to the assessment themselves, the reports, manuals, electronic resource, website and feedback sessions.

Assessment

The Evaluators (Evaluator 1, 2 April 2010; Evaluator 2, 17 February 2010) noted a ceiling effect on some of the subtests (Ideas about Maths and Counting) in the assessment that were repeated at the follow-up assessment. Both evaluators advocated for the inclusion of more advanced items in the follow-up assessment to address this issue.

Report

Evaluator 2 (17 February 2010) found that the reports were well structured and well written. He felt that the representation of data in writing, graphics and numbers was a powerful method of presentation and increased accessibility. Both evaluators (Evaluator 1, 2 April 2010; Evaluator 2, 17 February 2010) noted that the results in the report were norm referenced as opposed to criterion referenced and suggested that criterion referencing be incorporated in later versions. This could be

accomplished by Rasch Modelling to benchmark the assessments. The benchmarks could then be linked to resources and skills still to be developed to further facilitate ease of use of the feedback.

Evaluator 2 (17 February 2010) also suggested that the introduction be reworked to be more interesting: “A more enthusing intro (with a focus on helping learners) maybe more motivating than an account of the project structure”. It was also suggested that the note that small differences in scores should be interpreted with caution, rather be placed with the first tables, so that it is not missed by the school users when consulting the report (Evaluator 2, 17 February 2010).

Manuals

The Evaluators (Evaluator 1, 2 April 2010; Evaluator 2, 17 February 2010) judged the manuals to be well structured and written. Both Evaluators cautioned though that the writing style might be too academic for teachers. Further evaluation of the manuals by educators was recommended.

Electronic Resource

The evaluators judged the electronic resource to be very useful. Evaluator 2 (17 February 2010) commented on the importance of this resource as it provides access without requiring internet connectivity. Both evaluators commented that the dataset on the CD/DVD was useful in facilitating additional analysis by schools. Both Evaluators (Evaluator1, 2 April 2010; Evaluator2, 17 February 2010) recommended that an ‘autorun’ function be included in the programming of the resource to support less computer-literate users. Evaluator 2 (17 February 2010) commented that this would make the resource more accessible and user-friendly.

Website

> The website is a powerful tool for up-to-date information and making further resources available. It is a welcome addition to the project, as it provides links, materials, games and printable materials. (Evaluator 2, 17 February 2010)

Evaluator 1 (2 April 2010) agreed that the website looked very professional, but suggested that the page “About SAMP” be updated to focus on SAMP as a monitoring system and not just as a CEA project.

Evaluator 2 (17 February 2010) noted that there were only white children represented in the picture graphic in the header of the website. The picture does actually include an Indian and African child, but the comment however indicates that a clearer, more representative image should be sought.

Feedback Session

Evaluator 2 (17 February 2010) expressed that the feedback sessions were important to facilitate use of the manuals and report. This would facilitate use and understanding if the language was too academic. He supported the handing out of certificates at the end of the twice-yearly feedback sessions. Evaluator 1 noted some of his personal observations of attending the 2009 follow-up feedback session:

> It was very nice to hear the representatives of the schools express their enthusiasm and satisfaction with the various components of the feedback system offered to the schools. It was very clear, the schools and the teachers participating in the project experienced the SAMP monitoring and feedback system as an important tool. (Evaluator 1, 2 April 2010)

The expert evaluator report data were supplemented by data from the school users gathered through the evaluation questionnaires. The questionnaire data are discussed in the following sections.

8.2.2.6 Results and Findings - Teachers and Management Questionnaires

The questionnaires begin with the evaluation of the various components and the overall feedback system, followed by a discussion of how the feedback is used and finally the effectiveness of the feedback system is evaluated. Results for both the teacher and management questionnaires are presented in a combined tabular format, so that the results can be compared. The results of the report questionnaire are summarised in Table 8.3 - Table 8.14. Not all comments could be re-produced, so selected comments, which represent the various views expressed, are presented in the tables.

Table 8.3: Evaluation Questionnaire Data Teacher and Management – Background information

Questions	Teachers			Management		
	English	Afrikaans	Sepedi	English	Afrikaans	Sepedi
N	6	3	5	7	2	4
Position (HoD/ Principal if applicable)	.	.	.	HoD: 6 Principal: 1	HoD: 2 Principal: 0	HoD: 1 Principal: 3
Number of years you have been a teacher	13	25	20	.	.	.
Number of years you have been an HoD (if applicable)	.	.	.	5	17	9
Number of years you have been a Principal (if applicable)	.	.	.	5	.	6
Number of years you have been involved in SAMP	4	4	5	4	6	5
Total number of learners in Grade 1 (whole school)	147	143	97	159	149	101
Number of learners in your class	34	30	40	.	.	.
Gender of teacher	Male: 0 Female: 6	Male: 0 Female: 3	Male: 0 Female: 5	Male: 1 Female: 2	Male: 0 Female: 2	Male: 0 Female: 4

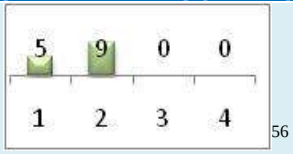
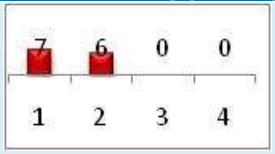
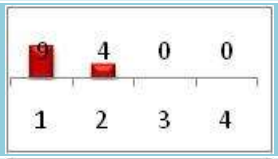
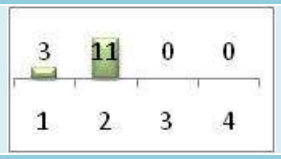
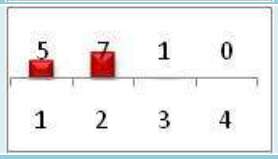
Pre-existing conditions to facilitate use

The various components (including the assessment, report, manual, electronic resource and website) are examined in term of their ability to facilitate use. This is followed by an examination of the overall conditions to facilitate use.

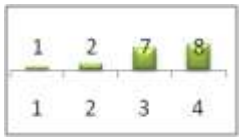
Assessments

The data on the Assessment component of the feedback system are represented in Table 8.4. Management found the assessments less disruptive than teachers did. Teachers who had children removed from their class for assessments experienced the disruption more directly. However, both the teachers and management indicated that the data were important enough to warrant any inconvenience caused by the assessments. The schools also expressed that they trusted that the results were a true reflection of learner abilities. All the participants indicated that their schools would continue participating in the feedback system.

Table 8.4: Frequencies of teacher and management views of the assessment component

Questions	Teachers (N=14)			Management (N=13)		
	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)
Although testing disrupts my class, the information we receive is worth the inconvenience.	1.6	2	 [56]	1.5	1	
My school will continue participating in the project if the assessments are conducted as currently.	-	-	-	1.3	1	
The results from the assessment reflect the learners skills and knowledge accurately.	1.8	2		1.7	2	
Any other comments .	• **Assesses all Grade 1 learners:** Would be very valuable if all the grade one learners can be assessed. Would like to know the skills and knowledge of all the learners in my class when assessed by SAMP. • **Good rapport with fieldworkers:** People are very nice and children feel comfortable at ease with them. • **Discrepancies between school and SAMP results:** At times there is a gap between the results i.e. 1,2,3,4 a learner get from us and from the university. • **Diagnostic value:** The assessments helps you the educators to identify areas of need i.e. where the learners experiences a problem.			• **Assesses all Grade 1 learners:** All learners should be tested, not only a certain number. • **Good rapport with fieldworkers:** They are very helpful and nice to work with sympathetic towards the children. • **Discrepancies between school and SAMP results:** There were some confusions with results from the assessment, as some outstanding learners who experienced problems received high results. • **Consider additional variables:** Please take into account our learners/social/ emotional backgrounds when doing comparisons.		

[56] The numbers below the X-axis: 1,2,3,4 represents the different categories on the scale, 1=Strongly Agree; 2=Agree; 3=Disagree; 4=Strongly Disagree. The numbers above the X-axis represent the frequency of responses for each category. In this example, eight respondents rated this aspect as 4, or Strongly Disagree. Please also note that, not all respondents answered all questions.



Under the general comments, the teachers and management both noted that they would appreciate it if all learners in the school could be assessed, and not only a limited number. Schools noted that the rapport between the children and fieldworkers was good. Some teachers mentioned the assessment's diagnostic value in identifying areas of need, while one principal asked that additional background and demographic variables be considered in the data analysis.

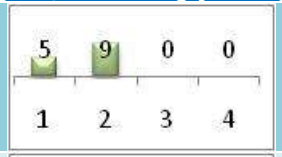
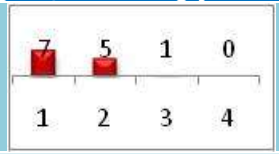
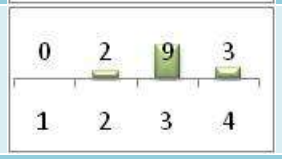
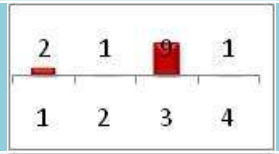
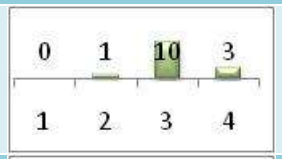
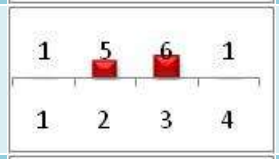
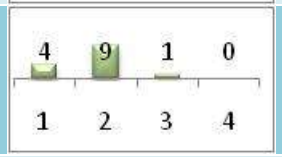
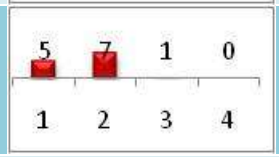
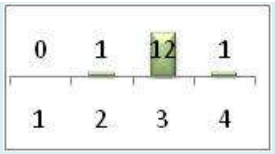
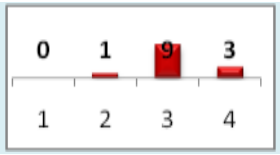
Some schools indicated that there were sometimes discrepancies between SAMP and school results. These discrepancies are to be expected as some learners may perform more poorly on the assessments if they struggled with the novel testing situation or to build rapport with the fieldworker. In other cases, learners may have performed better in the testing than in the class if they for instance have difficulty working in a group context or have unidentified hearing or vision problems. These comments however highlighted the importance of discussing why discrepancies may arise in the feedback sessions and providing information on this phenomenon in the instrument manuals.

Reports

The results of the evaluation of the report component of the feedback system (both the electronic and bound reports) are presented in Table 8.5. Reports contributed to the teaching roles of the teachers in the schools and the results represented in the reports were seen as trustworthy. The reports were seen as helpful for identifying areas of improvement and the turn-around time between assessment and reporting was judged appropriate to ensure that the data were still relevant for practise and planning.

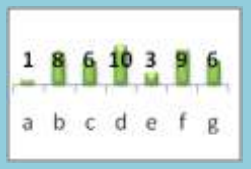
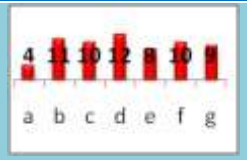
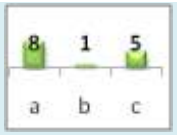
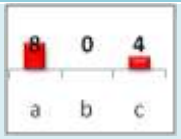
Although all the participants indicated, they could easily access the reports the management found this easier than the teachers. This is probably because reports are usually kept by the principals, HoDs and head teachers and teachers usually have to ask the management for access. It was interesting to see that the management did not see the reports as contributing to their leadership, only to teaching. It seems that the management participants did not see curriculum planning as part of their leadership role.

Table 8.5: Frequencies of teacher and management views of the report component (bound and electronic)

Questions	Teachers(N=14)			Management (N=13)			
	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)	Comments
I can easily get hold of my school's report if I need it.	1.6	2		1.5	1		I keep the records
The reports do not contribute to my teaching. (*)	3.1	3		.	.	-	-
The reports do not contribute to my leadership. (*)	-	-	-	2.7	3		Its not about leadership but teaching.
I often disagree with the information I find in the reports. (*)	3.1	3		2.5	3		Individual results. | Seldom
The reports help me to identify areas for improvement.	1.8	2		1.7	2		Interesting to see wether I assessed my children correctly.
The writing of reports takes too long, the information is no longer relevant when I receive it. (*)	3	3		3.2	3		No problem.

* Items stated in the negative in the questionnaire

(Table 8.5 – Continued)

Questions	Teachers	Management
Which parts of the report do you use the most and why? (a) introduction (b) overall school results in table form (c) overall school results per scale and subtest (d) individual learner results table (e) histograms of learner performance (f) identification of weak and strong learners (g) the summary and recommendations	 • **Monitoring of progress:** To know learners progress (Performance) • **Examining individual learner results**: … we definitely also look at individual learner performance. | …also for intervention and development. • **Identifying areas for improvement:** We want to improve the standard of Grade 1. | Helps me identify areas of need.| … for developmental purposes • **Grouping of learners:** To assist learners with similar challenges. • **Monitoring standards:** *Om hoe standaard te stel sodat prestasie van leerders goed vergelyk met skole in die ooste van Pretoria.*[57] |Compare with performance in class. | It is good to see what school made compared to other schools… • **Reporting needs:** Principal would like to see how our learners perform interesting to compare our findings to SAMP.	 • **Monitoring of progress:** Gives me an idea of learners progression from the initial tests. • **Identifying areas for improvement:** The manual and report help me to identify the area of development and strategies that we can use to help those who need assistance. | Gives background of school performance, able to identify areas which need improvement. • **Examining individual learner results**: Learners who need help-enrichment and comparing with other schools. • **Providing additional data for communication**: Helps me to comprehensively inform parents about learners needs and recommendations also when I fill in 450 support forms for the learner. • **Examining overall performance:** To get the overall performance of learners | All information received is absolutely important to understand performance of learners in school. • **Quality of reports**: Good/ comprehensive report.
Do you use the electronic report on the (a)DVD/CD (b)paper-based (c) both?	 • **CD/DVD**: I use DVD/CD report. It is straight to the point and user-friendly. • **Paper-based:** Paper-based: It is always at hand to use as reference. | Paper-based report is faster, when you want to use it in the classroom. (don't need a computer etc.) also use DVD/CD.	 • **Both**: Both paper-based and DVD/CD. I use the paper report to compare to previous years. • **Paper-based:** Paper-based report as it is easily accessible. Don't always have a computer available. | … grade educators meet in classroom, no computer available and saves time.

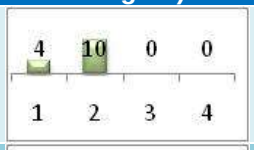
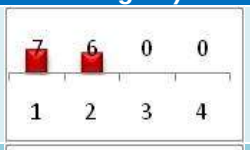
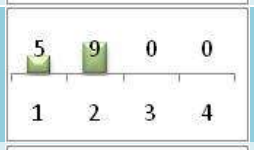
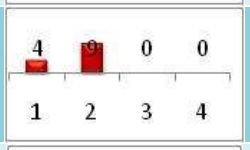
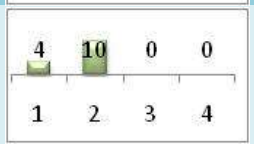
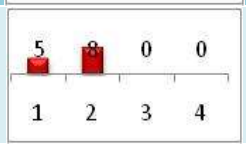
[57] English translation: *To set high standards, so that learners compare well to other schools in the east of Pretoria (more affluent part of Pretoria).*

Manuals

For teachers the individual learner results table, identification of weak and strong learners and overall school results in table form were the most utilised portions of the reports. The same three areas were of importance to the management participants, although the graphical representation of scale and sub-test results were consulted more often by managers than teachers. This seems to indicate that both teachers and management consult the overall school results from a curriculum perspective and the individual learner results to focus on individual cases. The management participants, however seemed to examine the overall school results in greater depth than the teachers who are mainly concerned with the overall relative performance of the school against others. The paper-based report was still noted as the preferred reporting style, but a third of the participants indicated that both the paper-based and electronic reports were utilised. The most significant limiting factor for use of the electronic reports was low levels of access to computers in the classroom.

The results of the items on the manual component of the feedback system are represented in Table 8.6. Both teachers and management respondents indicated that the manual was easily accessible in their particular school environment, although management participants again found it easier to access manuals as they usually kept the original reports and manuals with them. Manuals were judged useful in making sense of the data and rated as useful in planning and taking action to improve learning in the schools.

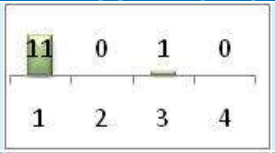
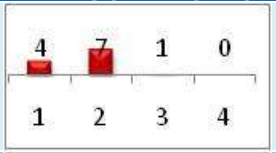
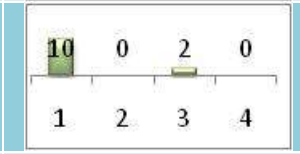
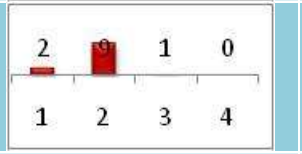
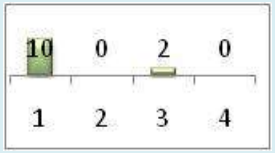
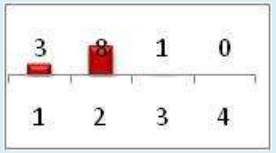
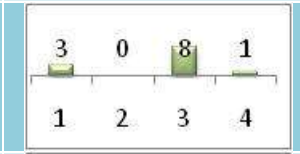
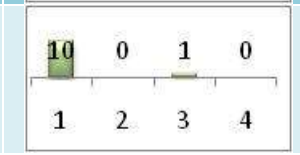
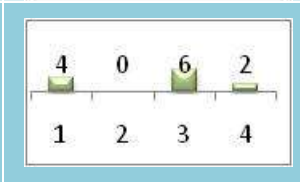
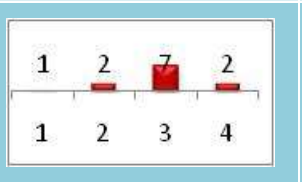
Table 8.6: Frequencies of teacher and management views of the manual component

Questions	Teachers (N=14)			Management (N=13)		
	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)
I can easily get hold of my school's manual if I need it.	1.7	2		1.5	1	
The information in the manual helps me to plan and take action to help learners to learn.	1.6	2		1.7	2	
The manual helps me to make sense of the information in the report.	1.7	2		1.6	2	

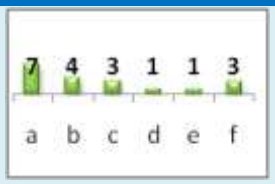
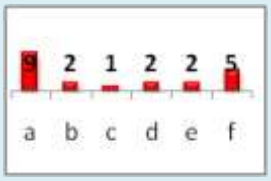
Electronic Resources

Table 8.7 illustrates the questionnaire results relating to the electronic resource provided at feedback sessions. Schools generally found the electronic resource accessible, although it was often kept by one person and not used by all staff. Unlike with the manuals and reports it would seem that one of the teachers usually kept the resource, which meant it was a bit harder for the management to access. Schools found that the resources were appropriate for their learners and the support materials were useful in helping to address areas of concern noted in the reports. Teachers found the resources more appropriate and useful than management, probably as resources are aimed at a classroom and not management level. In most cases, schools had however not passed on the resources to parents for home use and most schools had not employed the dataset for additional analysis. The printable materials were the preferred material on the electronic resource, although management staff indicated that they also accessed the electronic reports and manuals. Schools found the material child-friendly and appropriate. Respondents indicated that they did not have any suggestions for changes to the electronic resource, although some of the Sepedi resources were noted to be problematic in terms of spelling or translation.

Table 8.7: Frequencies of teacher and management views of the electronic resource component (CD/DVD)

Questions	Teachers (N=14)				Management (N=13)			
	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)	Representative Comments	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)	Representative Comments
I can easily get hold of my school's CD/DVD if I need it.	1.2	1		Did not receive it. | Do not have a DVD drive. | Not yet utilised	1.8	2		Have not used it I have found it somebody else had it.
The CD/DVD provides resources that are useful for teaching and learning.	1.3	1			1.9	2		
The resources provided on the CD/DVD help me to address the areas of concern mentioned in the reports.	1.3	1			1.8	2		
I have provided information or resources from the CD/DVD to parents or caregivers to use at home.	2.6	3		Not Yet. | Most parents do not have computers.	-	-	-	
The resources from the CD/DVD help learners in my class.	1.2	1		Not used as yet.	-	-	-	
I use the excel spreadsheet on the CD/DVD with my learners marks to do further analysis.	2.5	3		All the teachers do not have computers and are not computer literate!!	2.8	3		Not yet. | Do own assessment.

(Table 8.7 – Continued)

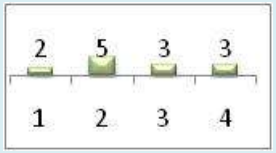
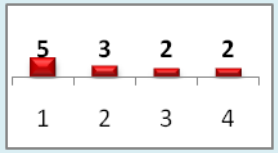
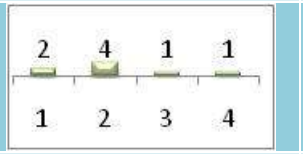
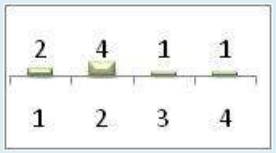
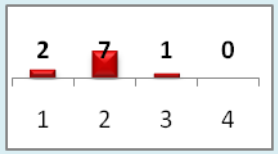
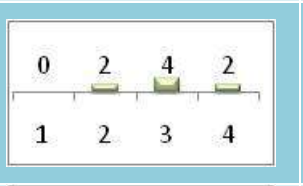
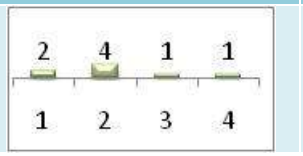
Questions	Teachers	Management
Which parts of the CD/DVD do you find the most useful? (a) printable materials for classroom use (b) printable materials for parents (c) electronic games (d) links to website resources (e) the electronic dataset (f) the electronic report and manual	 • **Classroom planning and use:** For lesson plan and remedial purposes. | The material can be used in the classroom. • **Distribution for home use:** They are easily accessible and to distribute to parents. • **Appropriate and attractive material:** *Leersaam en kindvriendelik.*[58] | The learners learn fast by seeing what to do.	 • **Appropriate and attractive material:** Attractive to learners. Learners enjoy working with them. • **Classroom planning and use:**| Able to use in the classroom. | Good ideas. • **User-friendly:** It is user friendly and provides valid activities and information.
Is there anything you would like to change about the CD/DVD?	No | Do not use the above-don't have time for that! We have too much school work !!!	No | Some spelling from CD (Pedi) they are not correctly spelled or translated.

Website

The questionnaire data relating to the website component of the feedback system are represented in Table 8.8. Accessibility to the website was problematic for many of the schools. In many cases, schools did not have internet access due to failure of the Gauteng online project of the DoE. In other cases, the cost of internet connectivity and limited computer-literacy were concerns. Schools judged the website content as appropriate and helpful to learners as well as to teachers for addressing areas of concern. Few schools had however provided details of the website to parents and caregivers.

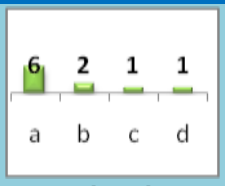
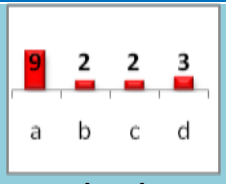
[58] English translation: *Educational and child-friendly.*

Table 8.8: Frequencies of teacher and management views of the website component

Questions	Teachers (N=14)				Management (N=13)			
	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)	Comments	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)	Comments
I am able to access the website.	2.5	2		Have not used it. | Not that computer literate. | Internet costly- person. Gauteng Online offline. | Do not have access to computer.	2.1	1		Have not used | Internet is very costly. Experiencing problems with Gauteng on line. Had to view using personal computer at home.
Resources provided on the website are appropriate for my class /school.	2.1	2			1.9	2		
The resources provided on the website helps me/the teachers to address the areas of concern mentioned in the reports.	2.1	2			2.0	2		Relevant.
I have provided information or resources from the website to parents or caregivers to use at home.	3.0	3		*Sal in 2de kwartaal deurgee aan ouers.*[59] | Not yet.	-	-	-	
The resources from the website help learners in my class.	2.1	2			-	-	-	

[59]English translation: *Will provide it to parents in the second term.*

(Table 8.8 – Continued)

Questions	Teachers	Management
Which parts of the website do you use the most? (a) printable materials for classroom use (b) printable materials for parents (c) electronic games (d) links to other website resources	 • **Classroom planning and use:** To use it together with the preparation - file & lesson planned for a day. | Adds to class resources • **Appropriate and attractive material:** *Kleurvol, kindvriendelik. Goeie hulpmiddel om les aanskoulik te maak.*[60] • **Difficulty with IT:** Internet not accessible. | It is difficult to use games as we don't have computers in the classrooms.	 • **Classroom planning and use:** • **Appropriate and attractive material:** They are on par with learner developmental level. They motivate learners, they enjoy working with them. |Material is easy and interesting for learners. • **Difficulty with IT:** Not using the Website as internet is not accessible. • **Integration with computer classes:** Good ideas. Games-computer centre.
Is the anything you would like to change about website or resources?	No | I am working on my computer skills and hope I will be able to access the Website soon. | Carry on we'll catch up.	No | Not at all.

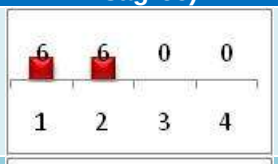
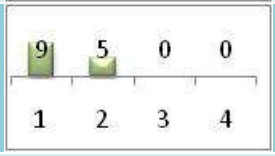
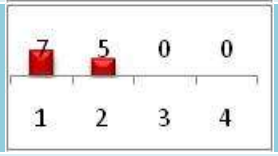
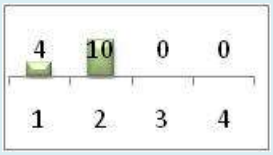
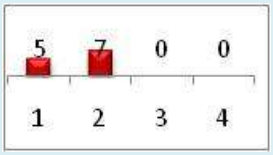
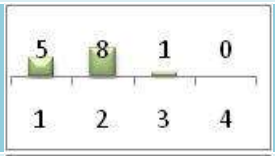
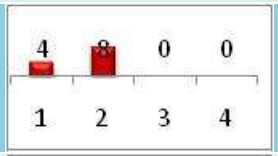
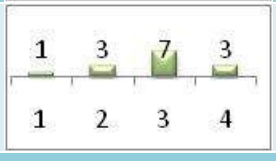
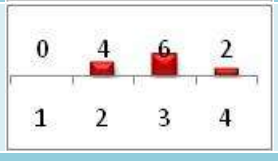
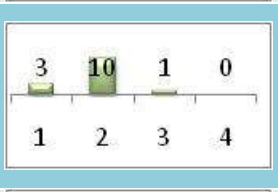
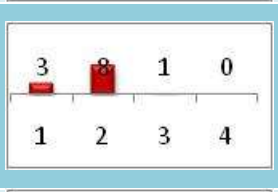
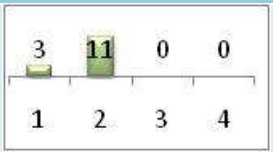
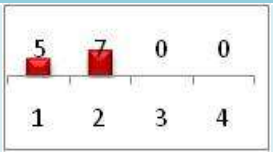
It seems that the management staff found it easier to access the website than teachers. This may be due to the fact that most schools have computers available for administration, computers for teaching and learning may not be available. The printable materials section on the website was the most used portion of the website as was the case with the electronic resource. Some schools noted, however, that the games could easily be integrated into existing computer classes. Schools commented that they would not like any changes to the website. Some teachers indicated that the schools were busy improving on their IT infrastructure and that they were working on improving their computer-literacy.

[60] English translation: *Colourful and child-friendly. Good support materials to make lessons more attractive.*

Feedback sessions

Table 8.9 relates to the data on the feedback session component of the questionnaire. The feedback sessions were deemed well-organised and presented. Schools experienced feedback sessions as helpful in interpreting the data and planning for improvement. Schools indicated that their questions and concerns were addressed appropriately in the feedback sessions. The participants indicated that it was possible for them to attend the sessions without too much inconvenience, although some principals noted that they sometimes sent a proxy in their stead.

Table 8.9: Frequencies of teacher and management views of the feedback session component

Questions	Teachers (N=14)			Management (N=13)		
	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)	Mean	Mode	Frequency distribution (1= Strongly Agree; 2= Agree; 3= Disagree; 4= Strongly Disagree)
The feedback sessions are well-organised.	1.4	1		1.5	1.5	
The feedback sessions are well-presented.	1.4	1		1.4	1	
My questions and concerns are addressed adequately during the feedback sessions	1.7	2		1.6	2	
The feedback sessions help me to use the reports and materials	1.7	2		1.7	2	
It is not possible for me to attend the feedback sessions without difficulty.(*)	2.9	3		2.8	3	
I enjoy having the opportunity to have discussions with teachers or principals from other schools at the sessions.	1.9	2		1.8	2	
My questions are addressed appropriately at the feedback sessions.	1.8	2		1.6	2	

(Table 8.9 – Continued)

Questions	Teachers	Management
Is there anything you would like to change about feedback sessions?	• **No changes:** Very good, organised. • **School visits:** How about visiting schools for feedback so that you get the in-depth feedback from teachers. • **Catering:** Yes. Cater for Halaal please.	• **No changes:** No, it to the point and very good. Organised. • **School visits:** If possible how about visiting individual schools for feedback so that you address each school report in isolation. • **Catering:** Yes. Please cater for Halaal.
What is your overall impression of the feedback sessions?	• **Generally positive evaluation:** Excellent | Very good. • **Worthwhile:** It is worth attending and helpful for my school's feed back. • **Professional:** *Uiters professioneel en opsommend*[61]. | Well-organised and well-presented. | Well-presented and informative • **Appropriate and concise:** To the point no waffle, knowledgeable people and relaxed atmosphere. • **Atmosphere and interaction:** Well planned, relaxed atmosphere. | Active participation and interaction with other educators at other schools i.e. learning from them.	• **Generally positive evaluation:** Good| Excellent • **Professional:** Well prepared. | Professionally organised • **Appropriate and concise:** It is super and short to the point not time consuming. |Information and well-organised. • **Atmosphere and interaction:** A very informative, well presented and interactive atmosphere.

* Items stated in the negative in the questionnaire

The feedback sessions were received very positively and perceived as being worthwhile. Both teachers and management indicated that the sessions were conducted professionally, considered appropriate to their needs, concise and presented in a relaxed atmosphere that encouraged collaboration. Most schools did not express a need for changes to the feedback sessions, although a suggestion for individualised feedback at each school was made. Some of the schools asked that the catering be expanded to include Halaal meals.

Overall conditions

A number of open-ended questions were used to gather information on the overall conditions required to facilitate use of the feedback system. The questions and selected comments according to themes are presented in Table 8.10.

[61] English translation: *Very professional and concise.*

Table 8.10: Open-ended questions on teacher and management views about overall conditions required for use of the feedback

Questions	Teachers	Management
How would you improve the feedback system?	• **Parental involvement:** Involve parents. • **No improvement required:** It's fine. | *Terugvoersisteem is uitstekend!* | No need to improve. • **Improve turnaround time:** Get feedback on time	• **Parental involvement:** Involving parents… • **No improvement required:** So far so good| No need to improve. • **Professional development:** … train educators.
What do you think of the facilitation of the SAMP feedback system?	• **Generally positive evaluation:** Very good…| Good |Baie goed![62] | Excellent • **Professional:** … professionally well done. | Very professional! • **Knowledgeable and dynamic:** Research team knowledgeable and dynamic • **Constructive**: It is fruitful and constructive. • **Logistically proficient:** Well-organised.	• **Generally positive evaluation:** It is well done. | Good | Very good | Fair • **Knowledgeable and dynamic:** Value-added research team is superb and dynamic, as they are constantly trying to acquire new developments and facilitate with excellence. • **Logistically proficient:** It is well-organised. • **Motivating:** Good and motivational

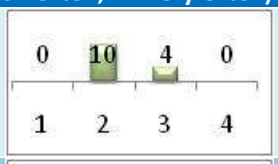
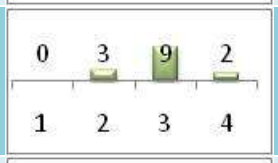
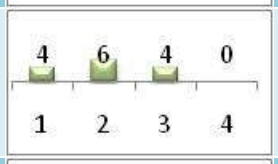
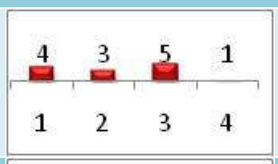
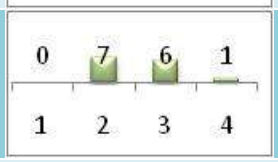
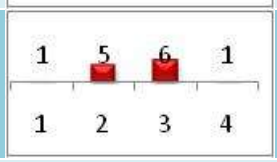
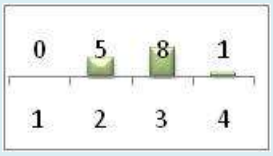
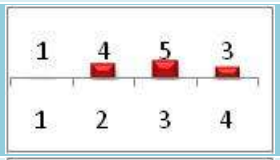
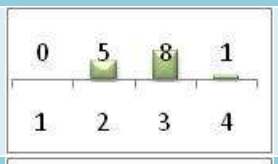
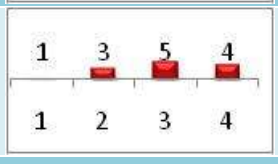
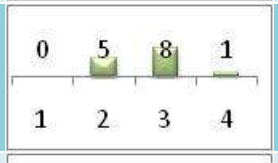
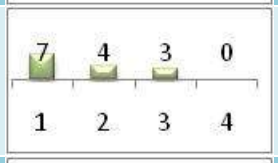
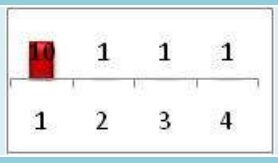
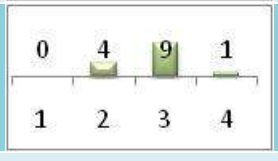
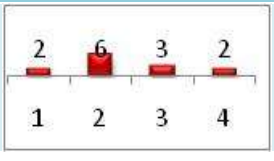
Teachers and managers generally felt that no improvement of the overall feedback system was required. Some teachers and management noted that involving parents would improve the feedback system, while one teacher asked for further improved turnaround time. Some of the management staff suggested additional professional development for teachers to use the feedback.

How schools use feedback

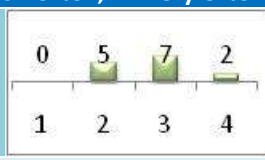
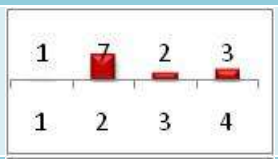
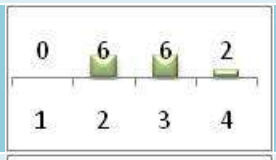
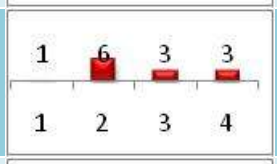
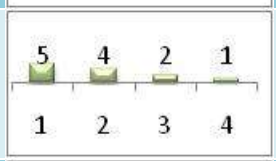
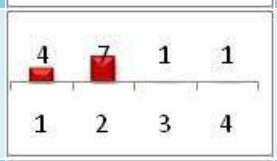
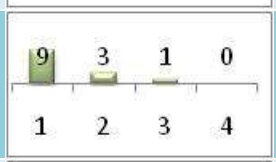
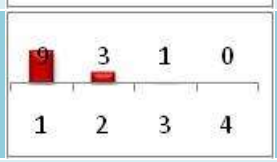
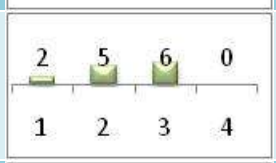
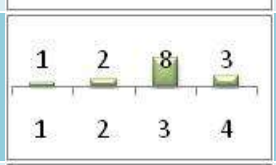
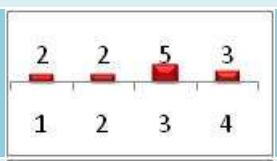
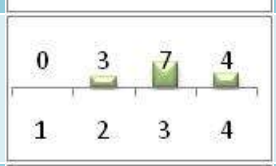
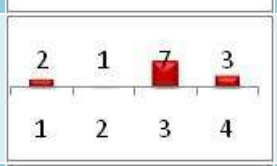
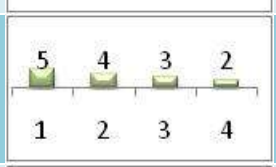
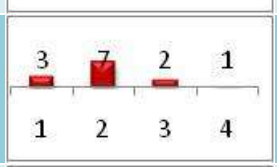
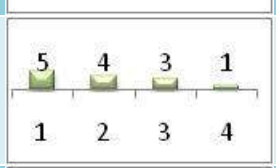
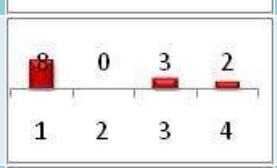
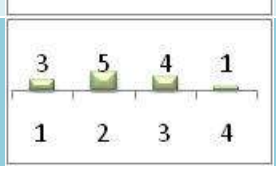
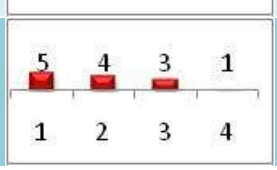
The different types of use made of the feedback system are noted in Table 8.11. All types of noted uses were to some extent employed by the schools. Teachers were less likely to use the data for the state mandated Internal Quality Management System (IQMS) and fundraising activities than management, as the IQMS forms part of the management function in schools. The teachers used the data most often for the purposes of self-directed learning, motivation of themselves and the learners, decision-making about learners, understanding aspects of their own teaching, professional development and comparing with their own standards of assessment. Management often used the feedback to inform school level policy, for planning purposes, to motivate teachers, and decision-making about learners, aspects which relate closely to the management function.

[62] English translation: *Very good.*

Table 8.11: Frequencies of teacher and management use of feedback

Questions	Teachers (N=14)			Management (N=13)		
	Mean	Mode	Frequency distribution (1= Never; 2= Sometimes; 3= Often; 4= Very Often)	Mean	Mode	Frequency distribution 1= Never; 2= Sometimes; 3= Often; 4= Very Often)
for instructional purposes.	2.3	2		.	.	
for self-directed learning (to help learners understand what they need to learn).	2.9	3		.	.	
to inform policy at the school-level.	2.0	2		2.2	3	
for planning purposes at the school.	2.6	2		2.5	3	
to motivate yourself and the learners	2.7	3		.	.	-
to motivate teachers	.	.	-	2.8	3	
for decision-making about learners.	2.7	3		2.9	3	
to understand aspects of my teaching	2.7	3		-	-	-
to assist in fund raising activities	1.7	1		1.5	1	
for my own professional development.	2.8	3		.	.	-
to further the professional development of the teachers.	.	.	-	2.4	2	

(Table 8.11 – Continued)

Questions	Teachers (N=14)			Management (N=13)		
	Mean	Mode	Frequency distribution (1= Never; 2= Sometimes; 3= Often; 4= Very Often)	Mean	Mode	Frequency distribution 1= Never; 2= Sometimes; 3= Often; 4= Very Often)
to compare with my own learner assessments	2.8	3		-	-	-
to compare with the teachers' standard of assessment.	-	-	-	2.5	2	
for internal monitoring of standards.	2.7	2.5		2.6	2	
as documentation for the Internal Quality Management System (IQMS)	1.9	1		1.9	2	
for discussions with : (a) district officials	1.4	1		1.4	1	
(b) the principal	2.3	3		-	-	-
(c) the foundation phase HoD	2.9	3		2.8	3	
(d) other Grade 1 teachers	3.1	3		2.9	3	
(e) teachers from other grades	2.1	1		2.1	2	
(f) pre-school	2.0	1		1.9	1	
(g) parents	2.2	2		2.0	1	

(Table 8.11 – Continued)

Questions	Teachers (N=14)			Management (N=13)		
	Mean	Mode	Frequency distribution (1= Never; 2= Sometimes; 3= Often; 4= Very Often)	Mean	Mode	Frequency distribution 1= Never; 2= Sometimes; 3= Often; 4= Very Often)
(h) Other, please specify: (n=1)	3.0	3.0	Specialists: school nurses, support educators, adopt- a- cop etc.	.	.	-

The feedback was regularly used to support communication, although rarely for discussions with district officials and the pre-school. The feedback was most often employed to support conversations with the principal, the foundation phase HoD and other Grade 1 teachers. Teachers used the feedback for conversations with parents more often than mangers, possibly as the teachers have more regular contact with parents. Management used the feedback to converse with teachers from other grades more often than the Grade 1 teachers did, as the feedback was used for a more holistic curriculum planning approach in these cases.

Further information about the use of feedback in the schools was collected through the general open-ended questions (see Table 8.12). Schools indicated that they felt they had contributed to the development of the feedback system. Both teachers and school management added that they contributed by taking responsibility for school improvement based on the data, sharing experiences and learning at the feedback sessions and recommending the project to other schools. The management users indicated that the schools also contributed through participation in evaluations and providing suggestions for improvement, as well as by being open and receptive to the feedback. From these answers it seems clear that the users felt a sense of ownership in the development process which increased the schools' receptiveness for the feedback.

Table 8.12: Open-ended questions on teacher and management views about use of the feedback

Questions	Teachers	Management
What do you think your contribution has been to the development of the feedback system? Please explain.	• **School improvement:** Improve teaching strategies. • **Sharing learning at feedback sessions:** To talk to other teachers about the standard and problem areas. | Sharing information as well as ideas. • **Recommending project:** Meer skole in sekere area te betrek. (names two schools in their area) ook nou deel van SAMP**.** [63] • **Logistical support:** To provide the learners for the tests.	• **School improvement:** Improving knowledge of teaching and learning. • **Sharing learning at feedback sessions:** Sharing of information during discussions ... |Sharing of experiences of learners difficulties. • **Recommending project:** Word of mouth| To talk to other teachers and schools. • **Attitude and receptiveness:** Open and positive about the feedback. • **Participating in evaluations and providing suggestions:** Giving feedback with experiences both positive and negative.
What other information do you take into consideration when you look at the feedback?	• **International results:** … international test results (through news reports) • **National results:** national… test results (through news reports) • **Own assessment:** … assessments in class| Mainly own assessments. • **Own experience: …** my own experience • **Additional variables:** Learners' background (Family and social issues) Parents' - level of education.| Own information regarding 5 year old turning 6 by the beginning of June. • **Data from other teachers:** … remedial class. • **Data from support practitioners:** O.T reports| I compare the reports with reports I have received from occupational therapists, speech therapists educational psychologists, etc. • **Comparative data:** We considered the level of our school with other school, the way of answering questions and speed when writing.| Own evaluation in relation to the university evaluation.	• **National results:** National Results… • **Own assessment:** Own assessments…| The information usually is in line with observations I do in the class about the learners and thus assures me of my decisions about learners performance… • **Additional variables:** In cases where the results are totally opposite to my observation, I would usually look at contextual factors that may have been an influence| Home circumstances, emotional state when tests are done, age of child. • **Data from other teachers:** …discussions with others. • **Data from support practitioners:** … occupational therapist reports, speech therapists reports. O.T. reports. • **DoE accountability data: …** 450 support information
What makes it difficult to use the feedback in your school?	• **Parental involvement:** Lack of parental involvement. • **IT infrastructure:** No electronic resources. • **Not all learners assessed:** Only 80 Grade 1's are tested. • **Computer-literacy:** I still struggle with the internet thing. • **Work load:** Time frame | Because of load of work … | Time.	• **IT infrastructure:** Lack of electronic resources. • **Not all learners assessed:** Because only 80 out 200 Grade 1's are tested. • **Work load:** Time. | Sometimes there are lot of school activities that need our attention we do not follow as we wish. | No time to get all educators together to use CD/DVD, as educators have to attend workshops, meeting, etc. • **No difficulties:** Nothing

[63] English translation: *More schools (names two schools in their area) now also part of SAMP.*

(Table 8.12 – Continued)

Questions	Teachers	Management
What helps you to use the feedback in your school?	***Characteristics of the feedback system:*** • **Overall integrated feedback system:** Everything • **Curriculum linkage:** Adapt the curriculum. • **Comparative component:** To compare our children. • **Recommendations:** Suggested activities in the manual. • **Differential and diagnostic data:** *Prestasies van leerders in elke leararea*.[64] • **Individual learner results:** Information on learners' performances. • **Structure and organisation of data:** Well organised makes it easier to present to others. • **Various presentations of data:** Graphs and percentages to show growth or areas of need. ***School-based actions:*** • **Meetings with other educators:** Meeting with other educators especially foundation phase educators. | It is helpful to sit as a panel of grade 1 educators and to discuss the results.	***Characteristics of the feedback system:*** • **Comparative component:** The comparison between our school and others is so clearly depicted …|… assessing the standard of the our school with other schools. • **Recommendations:** Suggested areas to work on-group and individual. • **Differential and diagnostic data:** Able to use the feedback to see which learners we can extend and which ones we need to remediate| For assessment, recording, reporting and identification. • **Various presentations of data:** Most of the words, sentences. Repeats: this helps in reinforcement and for those teachers needing support ***School-based actions:*** • **Motivation of educators:** Motivating educators • **School level planning:** The feedback is usually considered when we start at the beginning of the year, to identify areas that need more attention in Grade 1. • **Professional development:** I would then say it helps to inform the development of my teaching, methods and content. | To inform educators about new development e.g. CD/DVD. As well as suggested activities.
How do you decide which information in the feedback to act on?	• **Individual learner needs:** By looking at the learners' results and decide who needs support and referral and which area needs more attention | Problem areas and great achievers. | Areas waarin learders nie so good presteer het nie.[65] • **Meetings about data:** We have a face (sic) meeting with the principal and HoD to discuss areas of concern, then pick the most common ones. | Grade meeting • **Comparison to own data**: If something stands out, e.g. an exceptionally poor mark for a child, it stand out as something that needs to be acted on. The same for gifted child who needs more stimulation . • **Use all information:** All information Reading and spelling. • **Identified problem areas:** The problems are looked into.	• **Individual learner needs:** Look at weak children | Histograms of learner performance, individual learner results/ score. Helps to provide enrichment for high flyers and intervention and support for under achievers. • **Meetings about data:** Coming together foundation phase HoD and Grade 1 educators analysing the report and agreeing. • **Comparison to own data:** I look at the learners who have been referred and compare them with the report to see if they appear as support learners as well. • **Use all information:** Group and individual performance. • **Identified problem areas:** Areas that show a particular need for intervention e.g. maths… • **Relevance to context:** I look at the relevance of the information. | Taking all factors into consideration. • **Trend data for school:** Check report on results SAMP follow-up assessment … check … if we have improved or not. • **Comparison to other schools:** Check report on results SAMP follow-up assessment compare with other schools … • **Recommendation and conclusions:** Positives or negatives on the feedback are worth acting on. It is either to commended on good work or encourage improvements where necessary

[64] English translation: *Achievement of learners in every learning area.*
[65] English translation: *Areas in which the learners did not achieve so well.*

Although schools expressed that they found the data valid and trustworthy, they applied triangulation of data including use of own assessments, data from other educators, national and international assessments and data from other sources. Schools were incorporating the feedback data into their already existing knowledge base through a process of enlightenment prior to making decisions, planning and taking action. Although a variety of data were considered, the decision of which data to act on was made through collaborative processes of reasonability testing and feasibility testing. Reasonability testing took place by examining how the data compared to own and other assessment Once the data were deemed reasonable, feasibility took place by examining the relevance to the context and employing the identified problem areas and recommendations to see which areas should be addressed first with the school's resource pool of time, money and expertise.

Schools noted certain characteristics in the school environment which complicated use of the feedback including lack of parental involvement, poor ITC infrastructure and computer-literacy, as well as the heavy workloads limiting time to engage with the feedback. Schools also found the sampling of only a set number of learners for the monitoring a hindering factor.

The general question section re-affirmed the positive evaluation of the feedback system from the previous sections. The schools noted both characteristics of the feedback system and in the schools, which facilitated the use of the data:

- the overall congruence of the various elements of the feedback system,
- the diagnostic, differential nature of the data
- the comparative components and individual learner results
- the clear structure of the reporting materials with the multiple presentations of data.

For the teachers, the schools' characteristic that facilitates use was having conversations with other teachers. While for the management school level planning, motivation of teachers and professional development were seen as facilitating use of the feedback. Both teachers and management felt that the feedback system itself helped to establish these school characteristics through its structure, paradigm and support materials.

Effectiveness of the feedback system

The teacher and management ratings of the various components of the feedback system are illustrated in Table 8.13 and Table 8.14 respectively. Each component of the feedback system was rated in terms of appearance, clarity, need for the content, importance of the content, accessibility, ease of use and effectiveness.

The mode evaluation for every aspect of the feedback system was 4 or above. This shows a high level of satisfaction with each aspect of the feedback system for the evaluated quality criterion. However, it is apparent in both the teacher and manager evaluations that accessing and using the electronic resource and website is challenging. These resources also seem to be generally underutilised based on the data from the rest of the questionnaires. Poor IT infrastructure and low levels of computer-literacy seem to be contributing factors to this phenomenon. Schools however expressed the view that electronic aspects should be maintained as ITC infrastructure was improving and the teachers were working on improving their computer-literacy.

Table 8.13: Teacher rating of Feedback System Components – Histogram, Mean (Me) and Mode (Mo)

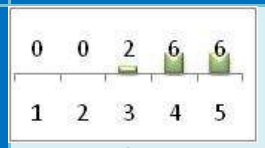
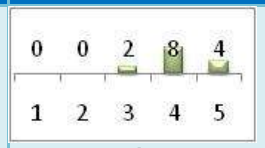
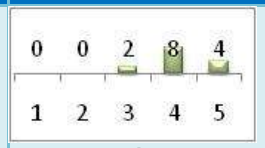
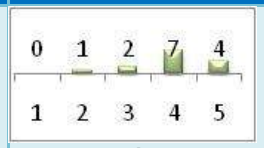
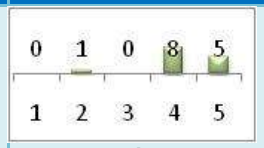
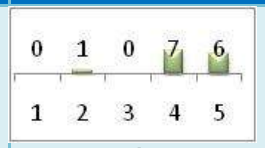
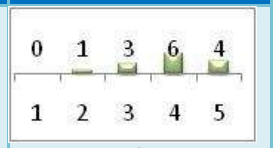
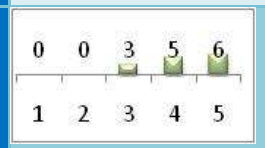
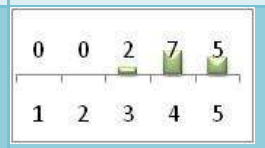
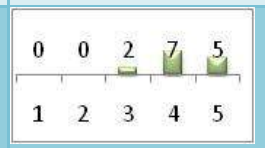
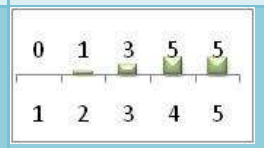
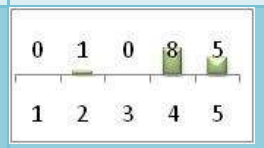
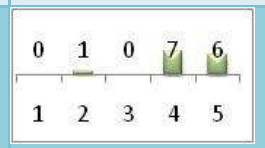
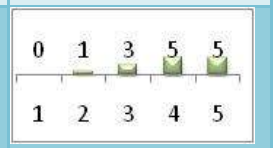
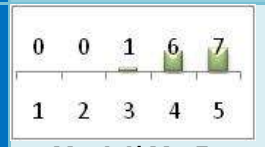
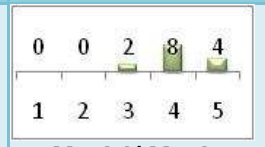
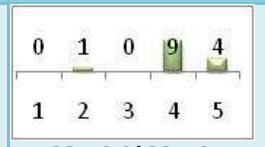
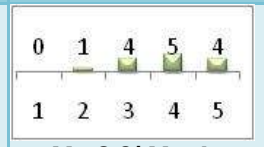
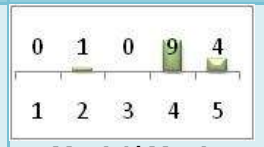
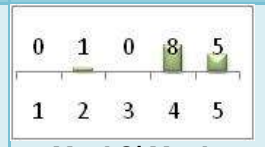
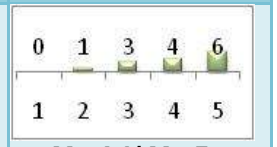
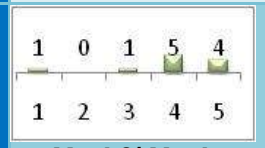
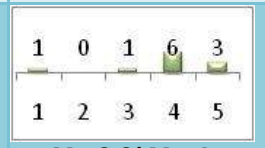
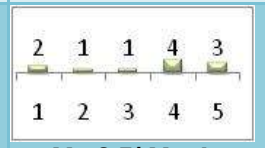
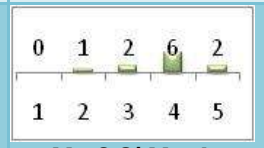
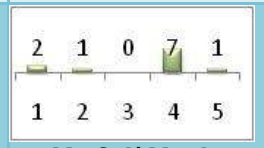
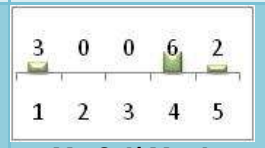
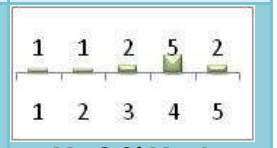
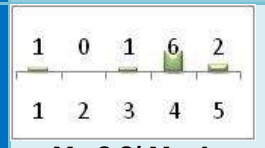
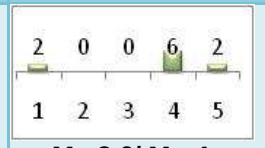
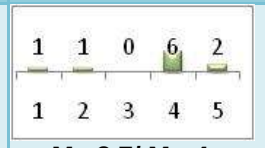
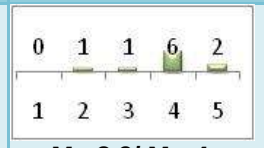
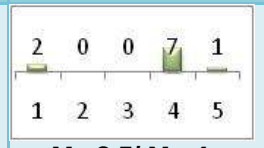
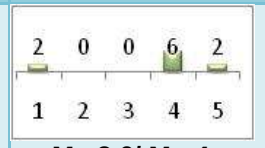
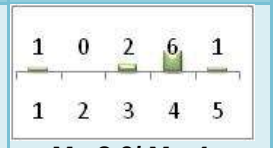
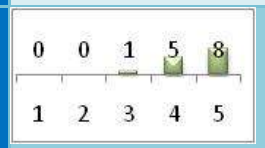
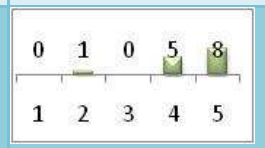
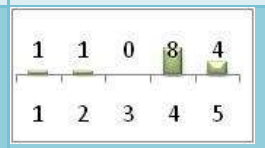
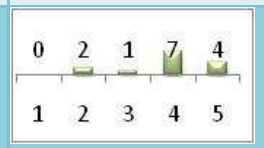
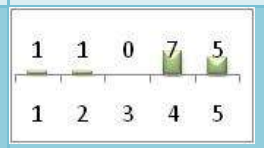
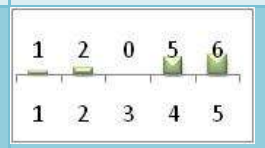
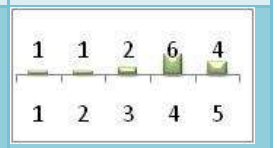
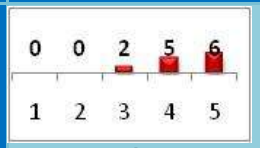
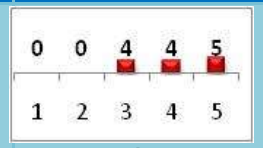
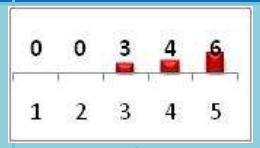
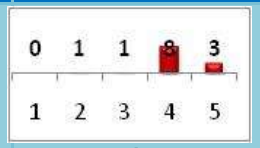
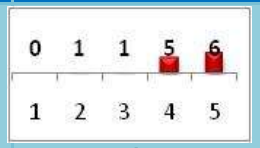
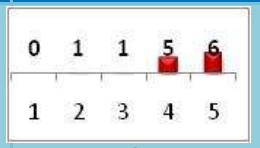
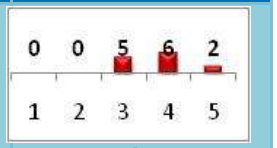
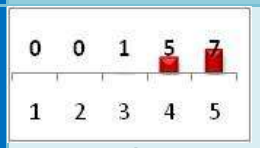
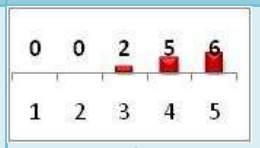
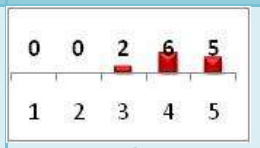
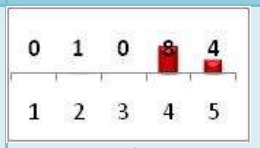
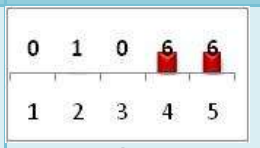
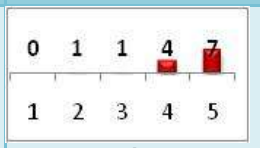
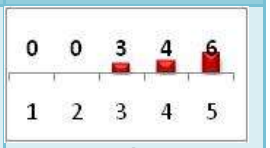
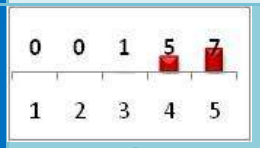
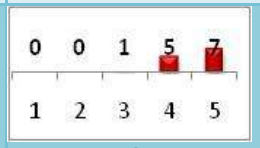
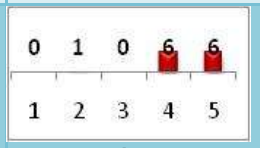
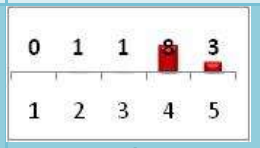
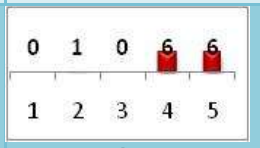
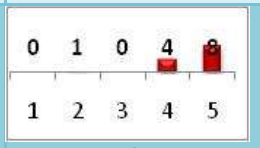
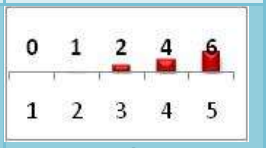
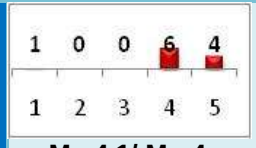
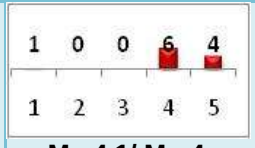
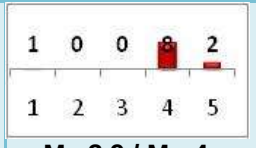
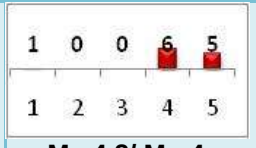
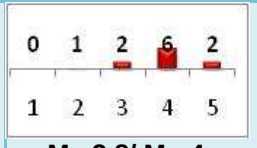
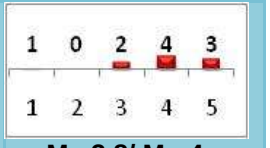
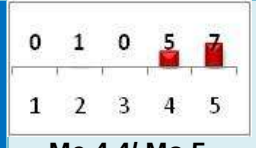
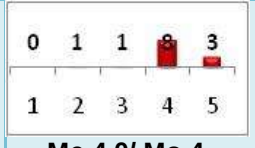
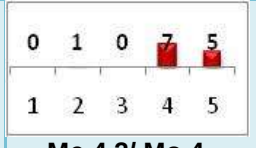
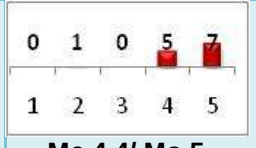
1=poor - 5=Excellent	Appearance	Clarity	Need for content	Importance of content	Accessibility	Ease of use	Effectiveness
Assessment	 Me 4.3/ Mo 4.5	 Me 4.1/ Mo 4	 Me 4.1/ Mo 4	 Me 4.0/ Mo 4	 Me 4.2/ Mo 4	 Me 4.3/ Mo 4	 Me 3.9/ Mo 4
Report	 Me 4.2/ Mo 5	 Me 4.2/ Mo 4	 Me 4.2/ Mo 4	 Me 4.0 / Mo 4.5	 Me 4.2/ Mo 4	 Me 4.3/ Mo 4	 Me 4.0/ Mo 4.5
Manual	 Me 4.4/ Mo 5	 Me 4.1/ Mo 4	 Me 4.1/ Mo 4	 Me 3.9/ Mo 4	 Me 4.1/ Mo 4	 Me 4.2/ Mo 4	 Me 4.1/ Mo 5
Resource CD/DVD	 Me 4.0/ Mo 4	 Me 3.9/ Mo 4	 Me 3.5/ Mo 4	 Me 3.8/ Mo 4	 Me 3.4/ Mo 4	 Me 3.4/ Mo 4	 Me 3.6/ Mo 4
Support Website	 Me 3.8/ Mo 4	 Me 3.6/ Mo 4	 Me 3.7/ Mo 4	 Me 3.9/ Mo 4	 Me 3.5/ Mo 4	 Me 3.6/ Mo 4	 Me 3.6/ Mo 4
Feedback session	 Me 4.5 / Mo 5	 Me 4.4/ Mo 5	 Me 3.9/ Mo 4	 Me 3.9/ Mo 4	 Me 4.0/ Mo 4	 Me 3.9/ Mo 5	 Me 3.8/ Mo 4

Table 8.14: Management rating of Feedback System Components - Histogram, Mean (Me) and Mode (Mo)

1=poor - 5=Excellent	Appearance	Clarity	Need for content	Importance of content	Accessibility	Ease of use	Effectiveness
Assessment	 Me 4.3/ Mo 5	 Me 4.1/ Mo 5	 Me 4.2/ Mo 5	 Me 4.0/ Mo 4	 Me 4.2/ Mo 5	 Me 4.2/ Mo 5	 Me 3.8/ Mo 4
Report	 Me 4.5/ Mo 5	 Me 4.3/ Mo 5	 Me 4.2/ Mo 4	 Me 4.2/ Mo 4	 Me 4.3/ Mo 4.5	 Me 4.3/ Mo 5	 Me 4.2/ Mo 5
Manual	 Me 4.5/ Mo 5	 Me 4.5/ Mo 5	 Me 4.3/ Mo 4.5	 Me 4.0/ Mo 4	 Me 4.3/ Mo 4.5	 Me 4.5/ Mo 5	 Me 4.2/ Mo 5
Resource CD/DVD	 Me 4.1/ Mo 4	 Me 4.1/ Mo 4	 Me 3.9/ Mo 4	 Me 3.8/ Mo 4	 Me 3.9 / Mo 4	 Me 4.2/ Mo 4	 Me 3.8/ Mo 4
Support Website	 Me 4.0/ Mo 4	 Me 4.0/ Mo 4	 Me 4.0/ Mo 4	 Me 3.8/ Mo 4	 Me 3.9/ Mo 4	 Me 3.9/ Mo 4	 Me 3.8/ Mo 4
Feedback session	 Me 4.4/ Mo 5	 Me 4.2/ Mo 4	 Me 4.2/ Mo 5	 Me 4.0/ Mo 4	 Me 4.2/ Mo 4	 Me 4.4/ Mo 5	 Me 4.2/ Mo 5

The quantitative ratings of the feedback system were supplemented with qualitative data from the open-ended question section of the questionnaires (see Table 8.15). A number of benefits of the feedback system for the school were noted. This included the system's diagnostic capabilities in identifying learners at risk or exceptional learners, the use of the feedback to monitor and improve standards, the motivational power of the feedback for schools and the value of the feedback for triangulation and monitoring internal assessment standards. The feedback was generally employed for planning and action in the schools and the managers indicated that the feedback helped with identifying professional development needs.

All schools indicated that they would recommend the feedback system to other schools. Reasons for recommendation included the overall effectiveness, inviting approach, diagnostic value, facilitation of monitoring standards, professional development opportunities, informing of teaching and planning and the quality of the data. Overall, the feedback system was evaluated positively with specific mention of the opportunities for professional development, dynamic and up to date nature of the system, high quality of the feedback and the usefulness in terms of monitoring and improving standards.

Table 8.15: Open-ended questions on teacher and management views about the effectiveness of the feedback system

Questions	Teachers	Management
What benefits has the SAMP feedback system had for you and your school, if any?	• **Identifying learners:** Also helps to identify learners who needs remedial. | It helps the Grade 1 teachers to identify and confirm problems. • **Monitoring and improving standards***: Om hoë onderrig- en onderwysstandaarde daar te stel.*.[66]| Improvement is showing/evident with our grade 1 learners. |…to know the trends or how all children in English schools are performing. • **Motivation:** *Skep die uitdaging om nou beter te presteer in vergelyking met ander deelnemende skole.*[67] • **Triangulation of own assessment:** Outside assessment of our learners' performance. • **Informing planning:** Planning to improve on areas where we are weak …. | Improve the planning of the curriculum.	• **Identifying learners:** ... to know where to focus our teaching on more. |… the grade one teachers, help them identify problem and or confirm their decisions • **Monitoring and improving standards***:* We are able to gauge our performances at a glance. | Helps us realise where our weaknesses are. • **Motivation:** It encourages us to work hard where we are lower and to keep up the standard where we are up. Different activities that stimulates educators and learners. | Has encouraged educators to work much harder at improving learners results. • **Triangulation of own assessment:** Bench marking, planning. • **Informing planning:** We know where or in which area we need to improve on. | We emphasize numeracy and literacy at grade R level to ensure reading ability and counting/numeracy skills at an early age. | Good idea of where we are going for the year. What we need to focus on. • **Professional development:** … educators are able to see their strength and weaknesses.
Would you recommend the SAMP feedback system to another school? Why?	**Yes, would recommend:** • **Diagnostic:** Helps in identification of learning problems. • **Monitoring standards:** It's a great way to see if your children are on standard. | … it is good to see comparisons in order to know if your school is on track. • **Professional development:** … a learning experience assessment. | Ja, verseker. Dra by tot verrykende denke! Nuutste onderrigmetodes en standaard van jou skool.[68] • **Informs teaching and planning**: It helps a lot in our daily teaching • **Data quality:** Provides vital information. .| Provides valid information. • **Invitational approach:.** It is non Judgemental… and free | It is constructive and educational in terms of educators' self|… it is well done, user-friendly and legible. • **Effective:** So that the school also can benefit and see the effectiveness of the programme.	**Yes, would recommend:** • **Diagnostic**: It helps to identify learners needs/ learning barriers • **Monitoring standards:** To see if their standard is on par with the other schools (national, local) | It is always important to get feedback and tests from an independent sector to inform the school about its status with regard to literacy and numeracy in comparison to the international standards • **Professional development**:. In terms of promoting early teaching of the relevant skills like at your ECD facility. • **Informs teaching and planning:** …informative. • **Data quality:** Comprehensive.| …. receiving valid information. • **Effective:** … so that they also benefit • **Professionalism:** Very professionally done. Person X's mannerisms and professional attitude makes it a pleasure and not a disruption.

[66] English translation: *To set high pedagogical standards.*
[67] English translation: *Creates the challenge to achieve better in comparison to other participating schools.*
[68] English translation: *Yes, definitely. Contributes to enriching and creative thoughts! Newest teaching methods and standards of your school.*

(Table 8.15 – Continued)

Questions	Teachers	Management
What is your overall impression of the SAMP feedback system?	• **Generally positive evaluation:** Excellent. It is really helpful to us.| Very good. | Love to attend it each year. | It is effective. • **Professional development:** Excellent and is of beneficiary (sic) for educators to improve their teaching. • **Dynamic and up to date**: Resourceful and quite up to date. • **High quality feedback:** Good and relevant. | Excellent and informative. • **Monitoring and improving standards:** Baie goeie riglyn om standaard van you skool se Graad Een leerders te bepaal en te vergelyk en te verbeter t.o.v. vorige jare.[69] | Helpful and it helps for self and school improvement or enrichment.	• **Generally positive evaluation:** Excellent. | Good work, keep it up. | Good | Very good • **Dynamic and up to date:** Informative, resourceful and on-going. • **High quality feedback:** Comprehensive, factual and detailed. Very relevant. | Informative and useful| A useful tool and very informative. Trends can be identified over a long period of time. | Good relevant and clear explanation • **Monitoring and improving standards: I**t provides us with detail on where to focus .

Several design guidelines were developed based on the evaluator reports and questionnaire data on Prototype IV. The design guidelines are discussed separately below.

8.2.2.7 Design Guidelines from the Evaluator Reports

1. ***Criterion referencing should supplement norm referencing*** to prevent educators from over-interpreting learner results.
2. ***Extension items should be added*** to the subtests where a ***ceiling effect*** is observed.
3. ***The difficulty level of reporting language*** should be evaluated and adjusted to make reports and manuals inviting and accessible.
4. ***User-friendliness of electronic resources*** should be facilitated as far as possible including an 'autorun' functionality, especially where users have low levels of computer-literacy.

8.2.2.8 Design Guidelines from the Final Evaluation Questionnaire

1. ***Data must be presented on several levels with diagnostic detail*** as reports are used on both individual learner, classroom and school level for planning. Data should for instance be presented graphically, textually,

[69] English translation: *Very good guidelines to determine the standard of your school's Grade One learners and compare and improve in comparison to previous years' results.*

tabulated and electronically to cater for different user preferences. The presentation of data at different levels of detail helps to provide data that can be applied to address different issues in the school environment, according to the school's needs.

2. ***Data must be presented in a non-judgemental, non-threatening way*** to encourage action as opposed to defensiveness.
    - ***Emphasis should be placed on both strengths and weaknesses*** of the feedback data. This ensures a balance between motivation and identifying areas for improvement.
    - This is particularly important at feedback sessions where ***a non-judgemental, collegial atmosphere must be engendered at feedback sessions***, to ***facilitate interaction and receptiveness.***
3. ***Fieldworker interaction with schools*** plays a large role in establishing the ***credibility*** of the assessments, data and feedback system as a whole.
    - ***Fieldworker training should be thorough and supervision on-going*** to ensure that fieldworkers interact professionally with schools and build good rapport with learners.
4. ***The quality of logistical arrangements and professionalism*** of the team contributes to ***trustworthiness*** of the feedback system***.***
    - ***Logistical arrangements must cause the minimum of disruption*** of classes and learning time.
5. Although difficulties with IT infrastructure and low levels of computer-literacy may hamper access to ***electronic resources***, the development and dissemination of such resources should not be ceased. Schools expressed a need for these resources, indicating that these resources were very useful and they would "*catch-up"* with their development of ITC resources and computer-literacy skills. The availability of such resources was seen as ***encouraging improvement of IT infrastructure and improvement of computer-literacy.***

- ***The use of electronic resources*** helped educators to stay current with developments in education and new resources. This supports ***professional development***.
- ***Resources incorporated must be both appropriate and inviting to learners*** to encourage and motivate them to use materials.
- ***Electronic educational games allow opportunities for cross-curricular work*** as for instance, literacy, numeracy and phonics games can be used in computer-literacy classes
- ***Websites*** provide the greatest opportunity for resources ***to remain current.***
- ***Website access*** allows for ***transfer of resources for home use.***

8.3 Conclusion

This chapter documents the Assessment Phase of the design research process. One complete cycle of design research took place during this phase, which examined the functioning of the elements of the feedback system, as well as the global functioning of the integrated system. The evaluation was achieved through expert evaluation reports as well as comprehensive questionnaires for teachers and management staff at schools. The overall evaluation of the feedback system was positive, with further design guidelines identified. In the next chapter, the data from the design research process are discussed along with the main design guidelines and recommendations for further research and design.

CHAPTER NINE
Conclusions and Recommendations

> When, after several hours, I came to myself again, I asked myself what it was that had so fascinated me. The answer is simple. The results were not presented as ready-made, but scientific curiosity was first aroused by presenting contrasting possibilities of conceiving the matter. Only then the attempt was made to clarify the issue by thorough argument. The intellectual honesty of the author makes us share the inner struggle in his mind. It is this which is the mark of the born teacher. Knowledge exists in two forms - lifeless, stored in books, and alive, in the consciousness of men. The second form of existence is after all the essential one; the first, indispensable as it may be, occupies only an inferior position. (Einstein, 1954, p. 80)

Providing quality education is a major concern for the South African government. Since 1995 the government has invested large portions of the GDP to improve education and rectify the educational inequalities of the apartheid era (National Treasury Republic of South Africa, 2005). Although there has been increased enrolment in RSA schools since 1995, the quality of education and poor learner performance remains a concern (Taylor, et al., 2003). It is often thought that increased funding will improve educational provision and consequently learner performance, but internationally increased expenditure in education has not delivered on this promise (Cassassus, 2001; Hayward & Hedge, 2005; Hattie, 2005).

Many believe that educational improvement can only be accomplished through feedback of educational data to schools (Coe, 2002). In RSA a wealth of educational data are generated through systemic evaluations, national and international assessments, whole school evaluations and classroom assessments. The availability of these data alone cannot improve learner performance, the data also needs to be returned to schools and appropriately employed by schools for planning, decision-making and action. Not all feedback translates into improvement, feedback often results in a negative effect (Fitz-Gibbon & Tymms, 2002; Kluger & DeNisi, 1996). This thesis examined how to bridge this gap between the availability of learner performance data and the feedback of the data in a manner that facilitates use of the data to inform planning and action in schools. (See Section 1.3 for a full discussion of the problem statement and rationale)

For the purpose of this study, an existing feedback system known as the South African Monitoring system for Primary schools (SAMP) was optimised. SAMP is currently facilitated externally by the Centre for Evaluation and Assessment (CEA). The system produces learner performance data for Grade 1 learners that are also aggregated to school level. The data from SAMP are employed to inform individual learner intervention, classroom practice, and school level planning and action. (For a full discussion of SAMP consult Chapter 2)

This final chapter concludes the research for this thesis. The research process is summarised in Section 9.1, followed by reflections on the conceptual framework (Section 9.2) and a summary of the research findings according to the research questions (Section 9.3). This is followed by reflections on the methodology (Section 9.4) and on the role of the researcher (Section 9.5). The conclusions and recommendations are represented in Sections 9.6 and 9.7 respectively.

9.1 Summary of Research Process

The aim of this study was to identify and understand the characteristics of an effective feedback system and the use thereof in order to design and optimise a system that facilitates the use of learner performance data in RSA within the school environment. The research question and sub-questions were as follow:

What are the characteristics of an effective feedback system and the use thereof in order to design an optimum feedback system to facilitate appropriate use of learner performance monitoring in primary schools in South Africa?

1. How can an existing learner performance monitoring system be appropriately adapted, contextualised and translated to the South African context?
2. What are the characteristics documented in literature of an optimal feedback system for use in school-based monitoring?
3. What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?
4. How do schools use feedback?
5. How effective is the feedback system in enhancing classroom practices, management and planning activities?

6. Which design guidelines for the development of an effective feedback intervention for school-based monitoring can be identified?

This research aimed not only to generate knowledge by describing the characteristics of an effective feedback system and developing design guidelines, but also to design and develop a well functioning feedback system. The main research question lends itself to a design research approach that aims to align research and utility (De Villiers, 2005; Van den Akker, 1999). The design research process is iterative and follows a cyclical pathway of development combining design, development and implementation, with formative evaluation to understand issues of application. In this study, each cycle of design research consists of the design and introduction of a version or prototype of the feedback system. This in turn is formatively evaluated leading to a further cycle of development with a new prototype.

The design research process for this thesis moved through 3 phases encompassing multiple design cycles:

- ***Preliminary Phase (one cycle):*** This phase focused on research sub-questions 1 and 2:

  1. How can an existing learner performance monitoring system be appropriately adapted, contextualised and translated to the South African context?
  2. What are the characteristics documented in literature of an optimal feedback system for use in school-based monitoring?

Sub-question 1 was addressed in Chapter 2 by examining how SAMP was adapted to the South African context, as part of the preparation for the Preliminary Phase. Sub-question 2 was addressed through the needs and context analysis, including a literature review (Chapter 3) and an exemplary case study of NZ (Chapter 5). The emphasis of this phase was to conceptualise the feedback system and define the design specifications. The evaluative foci for this phase were relevance and consistency (Section 5.2). Data for the NZ case study were collected through document analyses as well as interviews with school users, NZ Ministry of Education officials, asTTle development team members, professional developers and researchers using asTTle.

- ***Prototyping Phase (three cycles):*** Research sub-questions 3-4 were address in this phase:

  3. What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?
  4. How do schools use feedback?

This phase consisted of the iterative research cycles during which Prototypes I-III of the feedback system were developed, implemented and formatively evaluated (Chapters 6-7). The emphasis of the cycles shifted throughout the Prototyping Phase focussing first on how to establish conditions for use in Cycles 1 and 2 (Chapter 6) and then on how to transform these conditions for use into action in Cycle 3 (Chapter 7). Cycles 1-2 concentrated on research sub-question 3 with the evaluations concentrating on relevance, consistency and practicality. Cycle 3 examined research sub-question 4 with the evaluative foci of actual practicality and expected and actual efficacy.

For the first cycle, data were collected through expert evaluation reports and the Delphi technique with school users. The second cycle's data were generated through teacher and principal questionnaires. For the third cycle, questionnaires were again employed for teachers, principals and HoDs. This was supplemented with the examination of three schools' data-use processes through observations of school meetings, reflective journals for teachers and interviews with teachers, HoDs and principals.

- ***Assessment Phase (one cycle):*** This phase represents the semi-summative evaluation that examined the functioning of Prototype IV as an integrated system (Chapter 8). The phase therefore re-examined research sub-questions 2-4, but specifically focused on sub-question 5:

  5. How effective is the feedback system in enhancing classroom practices, management and planning activities?

The evaluative foci for this phase were therefore practicality and efficacy, both expected and actual. The semi-summative evaluation was conducted through questionnaires for both school management and teachers, as well as reports from

expert evaluators. This is the final phase of the design research process for this thesis, although further development may take place as part of the work of the CEA.

The design research approach for this study incorporates various combinations of qualitative and quantitative methodologies during each evaluation cycle of the prototypes. The different qualitative and qualitative procedures employed in this study are illustrated in Figure 9.1. The data from each evaluation served to develop design principles to inform the development of the next prototype which was again evaluated.(The overall research methodology is discussed in full in Chapter 4). In the next section, the conceptual framework is discussed as it framed this investigation.

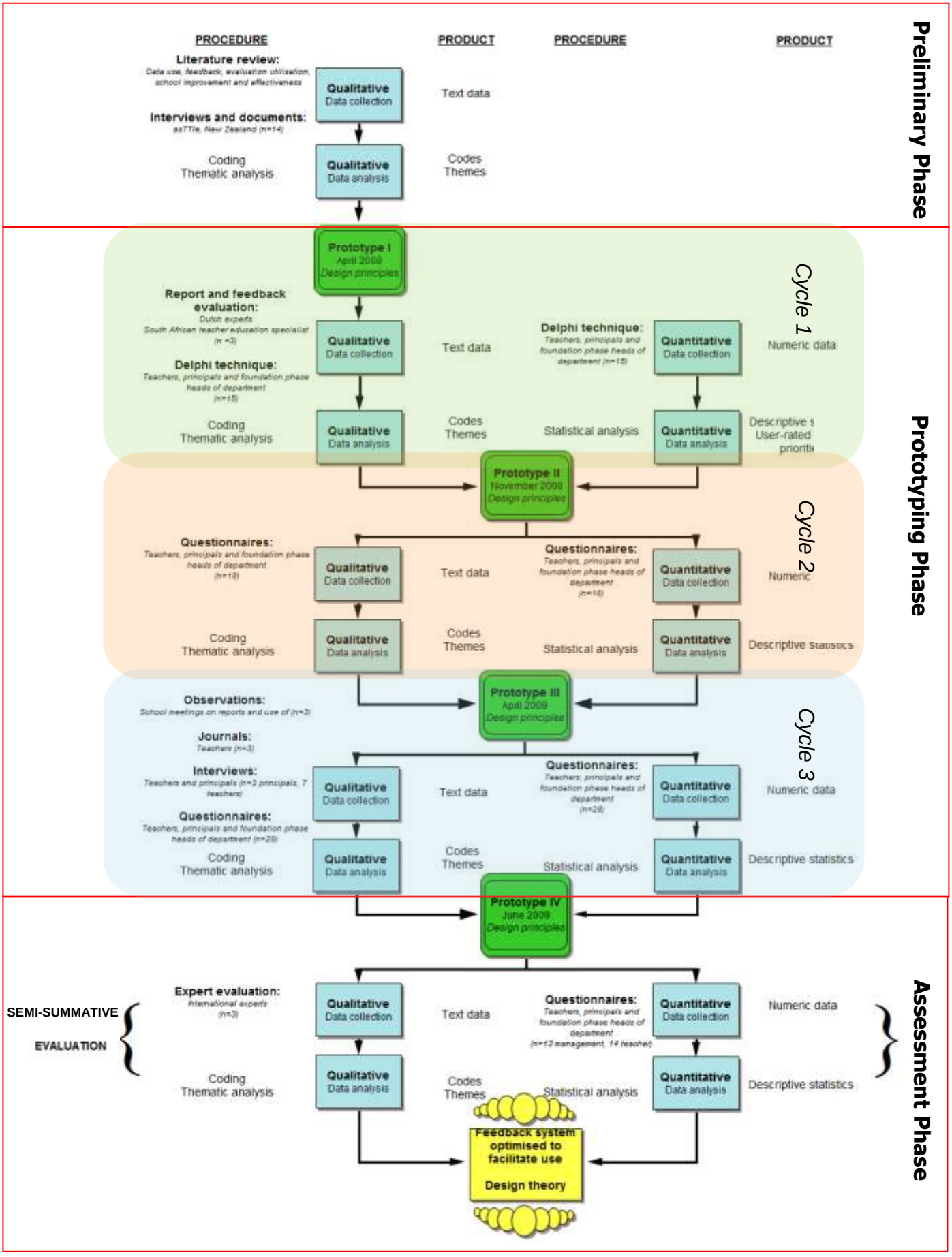


Figure 9.1: Research procedures

9.2 Reflections on the Conceptual Framework

The conceptual framework for use in this study (Figure 9.2) was developed from the literature on evaluation utilisation, SPFSs, feedback, school effectiveness and school improvement within a systems theory framework. How data are used in schools is often still a mystery and complicates the task of feedback facilitators in providing feedback that suits schools' data-use needs. Knowledge on data-use is mainly derived from organisational settings, psychology and the theoretical underpinnings of education as well as the field of evaluation (Coe, 2002; Kirkhart, 2000).

When examining use of feedback it is important to realise that use does not only pertain to direct use, but lies on a continuum from direct use to use purely for informational purposes without resulting in action (Alkin & Taut, 2003; King & Pechman, 1984; Weiss, 1981). Use in this research specifically refers to the process of applying the knowledge received toward a solution of a problem or the attainment of a predetermined goal (Love, 1985; Schildkamp & Kuiper, 2009). The application of the information may include direct use (instrumental use), delayed use or diffused conceptual use (Beyer, 1997; Estabrooks, 1999; Harnar & Preskill, 2007; Love, 1985).

Feedback interventions are often not used by schools, and if used in schools, often do not lead to improvement and under certain conditions even have an adverse effect (Kluger & DeNisi, 1996). The research focus has therefore shifted to facilitate the use of feedback to provide powerful opportunities for individual schools to analyse and improve quality of their education (Van Petegem & Vanhoof, 2005). The most salient feature influencing how receptive and responsive schools are to feedback seems to be the culture of the school (Salpeter, 2004). Schools that possess problem-solving capacity, innovation capacity and attitude, extra resources allocated to data-use in addition to tailored support and training, are more likely to show improvement due to feedback of data (Hulpia & Valcke, 2004). The school data-use culture is dynamic and can be altered through development, interaction with the feedback system and positive experiences of data-use (Schildkamp & Kuiper, 2009).

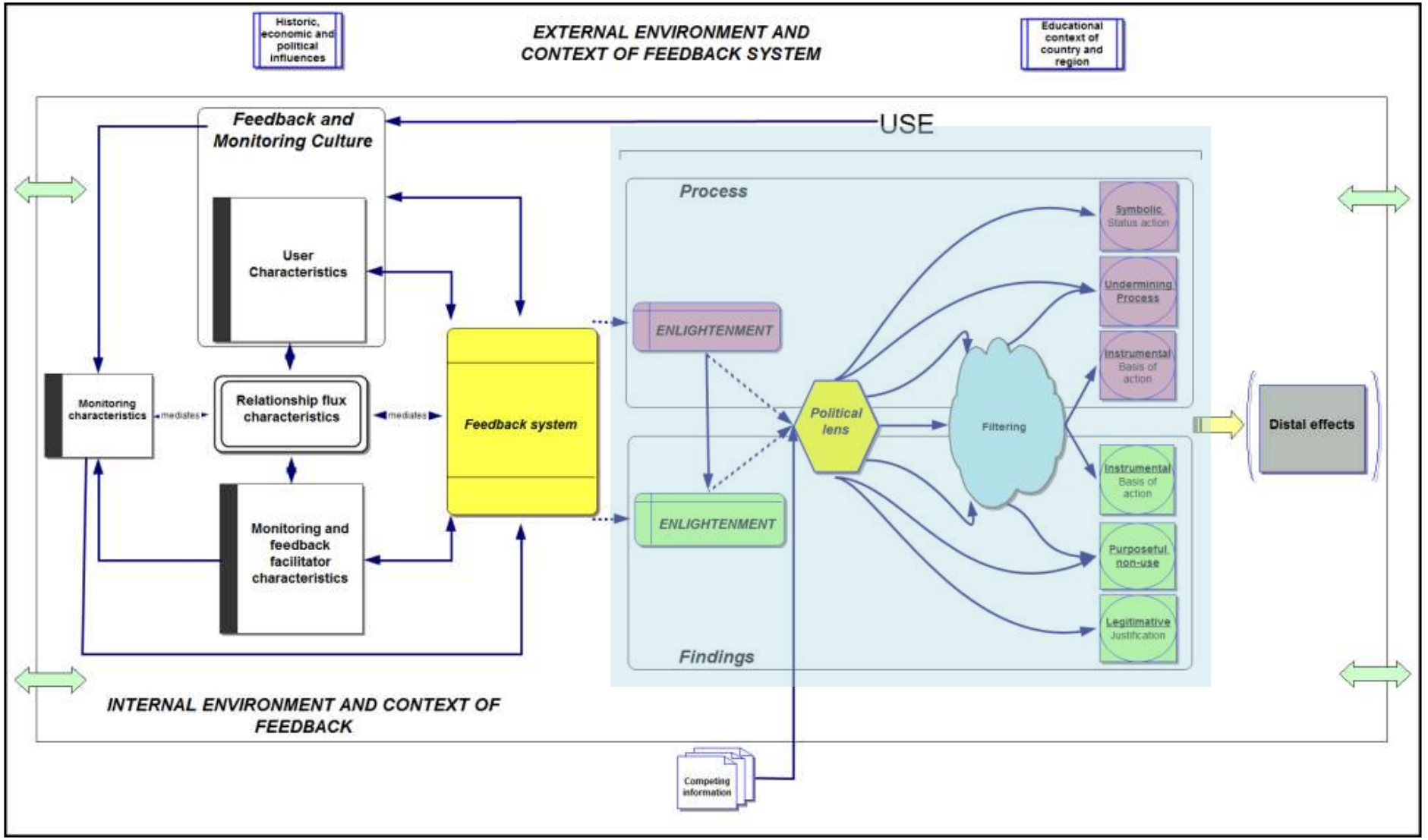


Figure 9.2: Original conceptual framework for use of monitoring feedback use

Data quality is important to facilitate use, data should also be relevant, reliable, valid, up-to-date, anonymous or confidential, have differential effectiveness and evoke positive reactivity. Presentation of data is also important, two-way communication and multiple modes of feedback presentation that considers user's level of data-literacy is more likely to evoke positive response (Bosker, et al., 2007; Brinko, 1993; Hulpia & Valcke, 2004; Schildkamp, et al., 2009)

A systems theory influenced the design of the conceptual framework. The framework acknowledges the interaction between various components in the feedback system that influence how feedback is received and how likely it is for the feedback to be used appropriately for decision-making and planning in the school environment. The original framework is shown in Figure 9.2.

The conceptual framework is a nested system acknowledging that any feedback is situated in the greater external education, monitoring and feedback environment in the country as well as in the immediate, internal context in which the feedback system and schools function. The use of feedback is influenced by the interaction of the feedback and monitoring culture of the schools, the characteristics of the monitoring system through which the data are generated, as well as by the characteristics of the feedback and monitoring facilitator. Combined with the characteristics and approach used in the feedback of the data these four groups of characteristics interact to determine the dynamic or flux characteristics (credibility, trust and sense of ownership) of the relationship between the users (schools, teachers and principals) and the monitoring and facilitator characteristics.

It is in this context that the data are provided through the feedback system which is characterised both by substantive and approach elements. Based on this two distinct types of use may take place, either process use based on what is learnt during the monitoring and feedback process or findings use which focuses on the data itself. In both cases of use, enlightenment or an enlargement of the body of knowledge first takes places. The information is then viewed through a political lens and then tested for reasonability and feasibility before it results in the different types of use. Every time use and participation in the feedback system takes place the monitoring and

feedback culture and experience of the users change and influence following cycles of feedback and use.

In view of these results, further adaptations were made to the final conceptual framework. The final conceptual framework for monitoring feedback use is shown in Figure 9.3. In the original conceptual framework, it was conceptualised that the feedback system only has an impact at the start of this process of use, providing data for enlightenment. The feedback system therefore was conceptualised as having limited influence on how the rest of the process of use takes place. Through the course of this research, it became clear that understanding of the data was seldom a concern, but rather, use was hindered when schools felt that they did not have resources to address the issues highlighted by the data. The feedback therefore failed the feasibility test. After deliberation with the research team, teachers, principals and HoDs it was determined that a multitude of education resources were freely available to support action and make acting on the data feasible, however three things prevented this process:

- Schools were often not aware of the availability of these resources.
- Resources, although freely available were often not collected in a central database, but scattered among resources from the DoE, the internet or at tertiary institutions or developed by specific schools and not shared with other schools. Internet resources were also available but the lack of internet access at schools meant that the resources were not accessible.
- Even if schools had the resources, they struggled to decide which resources to apply to the issues identified by the feedback.

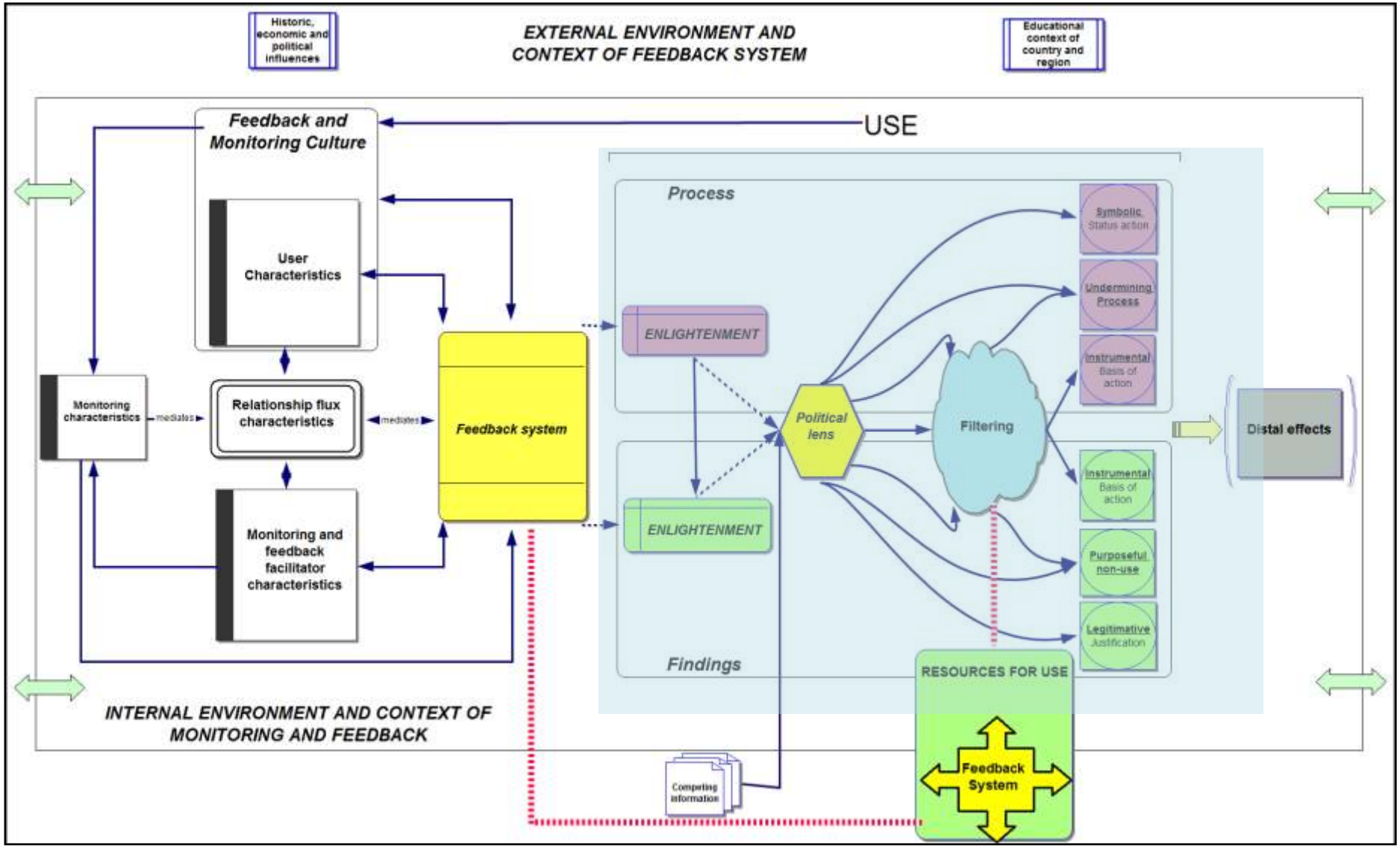


Figure 9.3: Conceptual framework for monitoring feedback use

An additional component was thus included in the conceptual framework. The feedback system was expanded to include a database of resources and provide links to resources to address certain learning areas and provide support materials to parents and caregivers. This was done by collecting the resources on a DVD/CD for schools so no internet access was required and publishing the resources on a web-page. All examples of interpreting and using data also now included a final section showing how the new resource database could be used to support improvement action. This shift meant that the feedback system's input in the process of use did not end with the provision of the data, but provided links to resources, which could positively influence whether or not the schools felt it was feasible to act on the data.

The contribution of this conceptualisation to the scientific body of knowledge on feedback systems for schools is that, although the context and feedback and monitoring culture of a school (Salpeter, 2004; Schildkamp & Kuiper, 2009) are paramount in how receptive and responsive schools are to feedback, these aspects can be influenced by the feedback system. Through the approach in development and interactions with schools, the culture of the school surrounding data-use can be influenced. At the same time providing links and a database of existing resources changes the awareness of the schools of the availability of resources that positively impact whether or not the data is acted upon.

Another shift in conceptualising of feedback systems is the shift in approach to facilitating understanding of the data. Often times the emphasis of support for understanding data built into feedback systems is on increasing the statistical literacy of users through for instance workshops, videos and materials that provide guidance on understanding the results (often very academic representations) as produced. Data complexity and lack of data-literacy are well documented as barriers to use of feedback for evidence-based practise (Black & Wiliam, 1998; Schildkamp, 2007; Schildkamp & Kuiper, 2009; Schildkamp & Teddlie, 2008; Visscher, 2002). This thesis advocates for an alternative approach where the existing processes and data-use styles of the users are examined in order to represent data in a manner

that is fairly intuitive and appropriate for the data-literacy level of the users. The data should still be presented in a statistically sound manner, but multiple data presentation forms can be used to reiterate reporting and make provision for different preferences of the users. Support to increase data-literacy should still be provided, but the focus should not only be statistical literacy. Rather, the support should be extended to understanding and interpretation of the data to be able to transform the data into planning, decision-making and action.

9.3 Summary of Results

The research results are presented in this section with specific reference to each research sub-question that was addressed. Reference is made to the chapters in which the full results for each sub-question can be found.

The overall research question for this thesis is:

> *What are the characteristics of an effective feedback system and the use thereof in designing an optimum feedback system to facilitate appropriate use of learner performance monitoring in primary schools in South Africa?"*

In order to address the overall research question, a number of specific sub-questions were explored. In this study, six specific questions were employed to answer the main research question. Each question and the relevant results are summarised separately.

9.3.1 Sub-question 1

> *How can an existing learner performance monitoring system be appropriately adapted, contextualised and translated to the South African context?*

Sub-question 1 was addressed in preparation for the Preliminary Phase. The question was addressed by documenting how the UK PIPS instrument was successfully translated, contextualised and adapted for use in RSA for Afrikaans, South African English and Sepedi (Chapter 2). The process incorporated statistical methods including an examination of reliability of

subtest and scales, Item facility and discrimination values and Rasch Analysis. The statistical data provided indications of scales, subtests and items that were of concern and required further examination. The data employed were the 2005 assessment data consisting of 417 learners. This was supplemented by teacher evaluations (n=6, two from each language group: Afrikaans, English and Sepedi) and the use of an expert panel (n=11, two educational psychologists, three teachers, two educational researchers and two subject experts involved in teacher education at a tertiary institution). The statistical analyses and evaluations also provided the means of re-examining the items after changes were implemented to ensure that the items were functioning appropriately and that valid inference could be made from the results. These processes were also supplemented by evaluations by two educational psychologists to map the overlap between the instruments and the curriculum. Substantial instrument-curriculum overlap was found with the Foundation Phase learning areas of Literacy and Numeracy. These curriculum links were included in the description of subtests in the instrument manual.

There are many approaches to adapt, contextualise and translate learner performance monitoring systems. However, it seems that any successful adaptation process cannot solely rely on statistical processes, but necessarily has to incorporate input from experts and users. This inclusion helps determine why certain items are problematic, increases contextual sensitivity, engenders trust and ownership and provides users with the opportunity to increase their statistical literacy.

9.3.2 Sub-question 2

> *What are the characteristics documented in literature of an optimal feedback system for use in school-based monitoring?*

The Preliminary Phase of the research design examined the characteristics documented in literature of an optimal feedback system (Chapter 4). A literature review was conducted and four international SPFS, namely SAM (Louisiana), asTTle (NZ), the CEM Suite (UK) and Zebo (Netherlands) were examined. The identified characteristics can be clustered into three groups:

technical characteristics; rollout and development characteristics and characteristics that alter the school environment.

Technical characteristics:

- ***Data*** must be relevant, reliable, valid, up to date, anonymous or confidential, allow for comparison, have differential effectiveness and evoke positive reactivity. (Angelle, 2004; Bosker, et al, 2007; Crooks 2002; Hattie, 2005; Hendriks, et al., 2001; Hulpia & Valcke, 2004;Tymms & Coe, 2003).
- ***Feedback*** should be thoughtful, reflective and focused to allow for exploration and understanding as well as two-way communication (Brinko, 1993; Bosker, et al, 2007; Hulpia & Valcke, 2004). Data is thus not just represented as fact, but presented in a way that invites interrogation and interaction. The feedback should also be provided in multiple manners employing for instance written reports, face-to-face feedback sessions and electronic representations (Brinko, 1993; Coe, 2002). The multiple representation styles accommodate various learning styles and preferences for users and provide multiple opportunities to interact with the data which may also reduce some users' anxieties about data-use and data-literacy. Short turn-around time for feedback is essential to ensure data are still relevant. This can be facilitated by ICT solutions that automate certain reporting functions (Angelle, 2004; Hattie, 2005; Hendriks, et al., 2001; Teddlie, et al., 2002).
- ***Reporting*** should consider user experience in data-use. This can be facilitated by employing different modes of presentation to accommodate user preferences (Brinko, 1993; Coe, 2002). The modes of presentation relates to how the data is presented (for instance graphically, textually or tabulated). Some users may prefer a specific data representation and using multiple presentations means that they may select how to view data and may even increase their data literacy through exposure to different data representations.

- ***Support*** should include tools to support school improvement-driven practices, (Angelle, 2004; Hendriks, et al., 2001; Teddlie, et al., 2002). These tools may include links to resources (either paper-based, electronic or web-based) but also includes making linkages between the data and curriculum apparent. Non-use of feedback is often due to a lack of knowledge of how to implement the feedback, not a lack of understanding of the data. Schools therefore do not know how to implement changes feasibly. Provision of tools to support use makes it easier for the feedback to pass the feasibility test and be used appropriately.

Rollout and development characteristics:

- ***User inclusion*** in the development processes through evaluations and consultation ensures contextual appropriateness and a sense of ownership (Hendriks, et al., 2001; Tymms & Coe, 2003). This in turn increases user familiarity with the feedback system and receptiveness to the feedback.
- ***Implementation through exemplars*** is recommended, e.g. starting with a small group to serve as living examples to other schools and teachers, followed by gradual dissemination of the approach (Black & Wiliam, 1998; Salpeter, 2004). Such an approach not only builds trust in the intervention's effectiveness in the context but also provides a network of users to discuss experiences and challenges in implementation.

Characteristics that alter the school environment:

- ***Data-literacy:*** Once again, the level of users' exposure to data-literacy can be improved by using multiple representations of the data and examples of how to use it, thus ensuring users are better equipped to work with data in the future. (Schildkamp & Kuiper, 2009; Van Petegem & Vanhoof, 2005). As data-literacy not only focuses on understanding data, but also the ability to use the data, support may also focus on improving the schools' problem solving and innovation capacity (Hulpia & Valcke, 2004)

- ***School culture*** directly influences how receptive and responsive schools are to feedback (Hulpia & Valcke, 2004; Salpeter, 2004; Schildkamp & Kuiper, 2009). Positive experiences of the feedback system and data-use in a non-threatening environment that is sensitive to user needs and input encourages a more data-friendly and responsive school culture (Hattie, 2005; Tymms & Albone, 2002; Schildkamp & Kuiper, 2009).

9.3.3 Sub-question 3

What pre-existing conditions need to be established in the feedback system to facilitate the use of the learner performance feedback system?

This research question was addressed during the Prototyping Phase, employing expert evaluations (n=3), the Delphi technique (n=15), a feedback session questionnaire (n=18) and a report questionnaire (n=28), discussed in full in Chapters 6-7. The exemplary case study of asTTle in NZ also provided additional information (Chapter 5). The research revealed that six components needed to be considered and attended to in order to have a successful feedback system.

1. ***Instruments*** must be designed, adapted, translated and contextualised in such a manner that they are not only valid and reliable, but also credible and trusted by the users. Participation of users in this process contributes to user sense of ownership of the feedback system and trust, which positively influence the monitoring and feedback culture of the school.
2. ***Reporting*** must be considered at all stages, i.e. during all parts of the feedback system the main goal of facilitating action and decision-making through reporting must be considered. Data must also be presented in a variety of formats, using different modes of feedback from written to face-to-face. This means that users may attend to feedback in the format with which they are most comfortable. In this way, users are not forced to adapt to a mode of feedback they are

uncomfortable with and can rather focus all their resources on understanding and using the data.

3. ***Support*** must be provided for ***understanding*** the data. Understanding the data is the first step in using the feedback appropriately. Such support should not only improve the users ability to understand the specific data, but also provides the opportunity to increase general data-literacy.

4. ***Support*** must also be provided for ***using data*** in schools, i.e. how to transform the data into action and decision-making in the schools. Once data is understood and problem areas identified for improvement, schools need several skills to use the feedback for decision-making, planning and improvement, for instance problem-solving, goal setting and monitoring skills. By modelling these skills in the feedback system (e.g. manuals, presentations, reports and electronic resources) and through examples, the users may improve these skills through process use. The provision of tools and links to resources also make it more feasible for schools to bridge the gap between understanding and effectively using the feedback, thus decreasing the barriers to use.

5. Throughout, attention must be paid to the ***relationships and communication between the schools and feedback system facilitators*** as these influence the trust, ownership and credibility of the feedback and the school's culture of data-use. Mutual respect is essential and can be shown through clear, concise meaningful communication, careful attention to logistical matters and training of the research and fieldwork teams.

6. These five components must all show a congruency that supports an underlying philosophy of assessment for learning and use of data for evidence-based practise as opposed to pure statistical understanding. This will support a ***paradigm shift*** in users, which may through process use be transferred to other aspects of teaching and management in the school environment.

(See Chapter 5-8) for a full discussion of these various elements)

9.3.4 Sub-question 4

How do schools use feedback?

This question was addressed in Cycle 3 of the Prototyping Phase by observing how three schools used the feedback from SAMP (Chapter 7). Observations, interviews and semi-structured journals were employed to examine the actual and expressed use of the feedback. This data were supplemented with data from the Assessment Phase questionnaires (management n=13, teachers n=14) in Chapter 8. The three schools investigated showed different levels of data-use sophistication, but all three schools used the data as the basis for planning, decision-making and action. All three schools used the data for several purposes: to support conversations with multiple stakeholders (amongst others HoDs, principals, other teachers, parents and departmental representative); to guide professional development; for curriculum development and planning; to meeting accountability demands; for triangulation; for further analyses, goal setting and monitoring

Three distinct approaches to data-use that appear to be effective emerged: *Team*, *Top-down* and *Cascade* (For a full discussion see Chapter 7). The most appropriate and effective model for data-use seemed to depend on the culture of the school, school leadership approach, level of teacher development, context and current level of functioning of the school. It was clear that a more advanced, sophisticated approach to data-use may be disheartening and inappropriate in certain contexts.

Once it was determined how schools use the data from the feedback system, it was necessary to examine how effective the SAMP feedback system was in enhancing classroom practices, management and planning. This was addressed through answering sub-question 5.

can be clustered according to guidelines for: instruments, reporting, support to understand data, support to use data, school relationship management and support for paradigm shift. The detailed design guidelines can be found in Chapters 5-8 and are only summarised here.

Instruments

The data generated to be provided through the feedback system must be shown to be reliable and allow for valid inferences for a feedback system to be effective. User involvement in adaptation, translation, contextualisation, development or evaluation of instruments is strongly advised to encourage trust, credibility and sense of ownership. Data must be differential to have diagnostic value and be curriculum-aligned to facilitate using data for decision-making and planning.

Reporting

User preferences should be accommodated through different modes of feedback (for example face-to-face, written and electronic) and incorporate various data representations (for example tables, graphs and text). Data must have comparative elements and should be confidential. Reporting should include both positive and negative feedback and include interpretations and recommendations to support evidence-based improvement practise. Employing these guidelines decreases the demands on the statistical-literacy of the users, provides opportunities for users to improve their data-literacy and increases the school's receptiveness and responsiveness to the feedback.

Support to understand the data

Incorporating various formats of data representation facilitates understanding of data, but should be accompanied by explanations, examples and support material. Support must be provided in a variety of manners such as written manuals, electronic support, web support and live interaction through feedback sessions and telephonic support, some of which should be available around the clock. This type of support provides users with the opportunity to select the most appropriate support for them and accommodates users with different levels of data-literacy skills.

Support to use the data

Once users understand the data, the next step is to use the data for improvement action in the schools. This can be supported by including interpretations, recommendations and links to tools for action in the feedback. This type of support should again be represented in a variety of modes with some support being available around the clock, e.g. printed materials, electronic resources and web-based support.

School relationship management

Every interaction with the schools provides an opportunity to alter perception and increase the receptiveness of the school users to the feedback. The quality of interactions is more important than the frequency of interactions. Communications should be clear, concise, respectful and encourage two-way communication that values user input. Fieldworker training is an essential component of school relationship management. A record keeping system of communications is essential to prevent duplication of communication by other team members. Professional execution of logistical matters provides an opportunity to show respect for users and improve the relationship flux characteristics between the feedback facilitator and users.

Support for paradigm shift

A learner performance feedback system can be a powerful tool to facilitate paradigm shifts. In this case, the feedback system aimed to entrench certain concepts with the users: use of data for evidence-based practise; the need for differential teaching; assessment for learning as opposed to assessment of learning; greater understanding of the curriculum. Whatever the underlying paradigm of a feedback system, all the elements of the feedback system should embody this and be congruent with the other elements. For example, modelling the approach to interpretation, planning and action based on the data, can be a powerful tool to embody the paradigm and support process use of these skills in the schools.

9.4 Reflections on Methodology

The use of a design research approach was highly effective for the design and adaptation of the feedback system. Feedback systems must be contextually appropriate and suit the preferences of the users while reporting the data in a reliable manner that provides the opportunity for valid inferences and achieves consequential validity through evidence-based practise in the schools. The design research approach allowed the opportunity to design, implement and evaluate various prototypes, which slowly started to approximate the ideal for the specific context. Design research includes representatives of the target users in designing the interventions. This meant that users could feel a greater sense of ownership of the feedback system, making them more receptive and responsive to the data. The design research approach had a greater cost benefit than carrying out a traditional design experiment to evaluate the existing feedback system that would merely have shown the limited effectiveness of the original system. Such an experiment would then have to have been followed up with further costly design to address the problems.

The design research approach also allows for the use of mixed methods as employed in the study. The combination of methods meant that the different weaknesses inherent in qualitative and quantitative methods could be supplemented with the different strengths of the methods, leading to more robust results. For this study the quantitative methods provided statistical data that were essential to evaluate and rate the effectiveness of the system and design the instruments. The qualitative methods provided insight into the user experiences and processes of use, which were essential to study how the data were used and study the feedback and evaluation cultures in the schools. The mixed methods approach also ensured multiple points of data source triangulation. Design research also requires input from both experts and users, which provided insight from both an academic and contextual level. The design research philosophy of viewing the users as true partners and collaborators in the design process was congruent with the collaborative approach used in the feedback system. The participation in the design research process also afforded users the opportunity for process learning about evaluation processes and furthered data-literacy of the participants.

A design research approach is very labour intensive and cannot be conducted in a short time frame. This has the benefit of prolonged engagement, but has cost implications. In this research four different prototypes of the feedback system were developed over a period of four years. This research concludes at this point and it seems that all the necessary evaluative criteria have been fulfilled. However, a cost-benefit analysis will still have to be conducted to determine if further development of the system is required.

Small sample size is a characteristic of design research (Plomp, 2009). In this research this can be seen in particular in Cycle 3 where the various approaches to use were examined. Only three schools could be followed. It is highly possible that additional approaches to effective use could be identified by studying a greater sample. The three schools were also specifically chosen as they were using the feedback effectively. Studying poor use, or non-use of data would provide the opportunity to gain a greater understanding of use, especially through comparison of approaches to effective and in-effective use or non-use of data. Unfortunately, this was beyond the scope of this thesis.

Some known design research dilemmas include the complication of working in a real world setting, where you as researchers may be a cultural stranger and need to be adaptable (Plomp, 2009). In this research, the researcher was a cultural stranger in the school environment. This was however overcome by collaboration and forming mutually beneficial relationships with the school users. The design research must be adaptable as each cycle is based on the previous. In a real-world setting changes in the research design are inevitable. Some consistency is essential though and this was ensured by having an explicit conceptual framework (Section 3.5) and congruent study design (Chapter 4), employing both deductive and inductive reasoning as well as rich descriptions.

One of the larger dilemmas in this study relates to the dual role as evaluator and facilitator of the feedback system, a characteristic of design research. Several measures suggested by Plomp (2009) were taken to compensate for this potential conflict of interests:

- ***Opening the research to professional scrutiny to people outside the project,*** including the three expert evaluators, research and educational psychologists and colleagues acting as critical friends.

- ***Shifting from the 'creative designer' perspective to the 'critical researcher' perspective as the research progresses:*** In the Preliminary Phase and early Prototyping Phase methods were used that explored and generated ideas, e.g. the NZ case study and Delphi technique. Later, methods such as the questionnaires and final expert evaluation reports were aimed at a more critical research perspective.

- ***Having a good quality of research design:*** Throughout the various design research phases, each cycle was given equal importance and attention. This helped establish a *strong chain of reasoning* (Krathwohl, 1998). Triangulation of method (qualitative and quantitative) and sources (users and experts) was employed to *explore empirically* the consistency, relevance, practicality and effectiveness of the intervention. Process notes, memos (*Atlas.ti)*, reflections (research diary) and member checks (interviews) were *systematically documented* to provide an audit trail. Teachers, HoDs and principals as well as the expert evaluators and colleagues acted as *critical friends*. The *quality of the data and instruments* was ensured through the methods discussed in Chapter 2 as well as trialling of evaluation instruments. The dual role of the researcher is discussed in depth in the next section.

9.5 Reflections on the Role of the Researcher

During this research, I held the roles of monitoring and feedback facilitator, as well as evaluator during the design research process. I explore these roles and implications for this research along with the role of self-reflexivity in managing these roles and the tensions in Section 4.3.1.

One of the main concerns with holding these multiple roles was how the act of evaluation would affect the attitudes of the feedback users towards the feedback system:

> Every time I have to ask the teachers, principals and HoDs to complete a questionnaire or participate in an interview or one of the other data collections. As a feedback facilitator I am always aware of the high administrative demands placed on teachers, principals and HoDs in schools in South Africa. Time to work with data is limited and I always try to make the feedback as user friendly as possible to ensure that the limited time is spent on deciding and planning how to act on the data as opposed to trying to make sense of it. It keeps feeling that the act of evaluation will have a quantum effect, negatively influencing the school's attitudes to the feedback system as a whole or taking time away from using the data in the schools. (Research diary 26/05/2008)

It soon became apparent that this was not necessarily the case:

> After the feedback session I realised that the schools actually appreciate the opportunity to give feedback. The key seems to be that the schools feel that their feedback is taken seriously and impacts how they receive feedback. If anything, the schools seem thankful for the opportunity and see it as a way of expressing their thanks for the feedback, which they find very useful. (Research diary 01/12/2008)

The evaluation process also increased the schools' sense of ownership of the data, gave them some insight into how evaluations can be conducted, and increased the users' familiarity with and understanding of the data and presentation formats. In effect, the evaluation process may have had a positive impact on the feedback and evaluation culture in the schools and receptiveness of the schools to use the feedback. It also seemed to have had a positive effect on the trust, credibility and mutual respect developed between the users and facilitator. This re-emphasised the need to try to represent the input of the users as well as possible and to reproduce their voices authentically in their evaluations and in the way I developed prototypes based on the feedback. In many cases, things that I was concerned about in prototypes were not the main priorities for schools and I had to pay attention to maintaining a balance between my own and evaluator inputs as well as user inputs.

The users' high levels of willingness to participate in the evaluations raised a new concern that the users were providing socially desirable responses, but

the multiple points of triangulation of sources showed a consistency of responses. Sensitivity about the demands placed on the users through the evaluation also provided extra impetus to design each evaluation very carefully and to ensure that evaluations were as clear as possible and took up the least amount of time, while still fully addressing the specific questions.

From a researcher facilitator point of view also being the evaluator meant that it was easier for me to be cognisant of the users' ownership and contribution to the feedback system. This made it easier to see the users as full partners in the feedback system and always ensure two-way communication.

From a very personal point of view I can be a bit controlling and the size of the study necessitated that I used multiple fieldworkers and a feedback facilitation team.

> I constantly have to try and moderate my own fear and need to control all communications with the school. If I gave in to it I would do all the phoning, faxes, letters and fieldwork myself and this is just impossible. One of the hardest things to do is to let go. The team has been amazing though, it seems that the training and guidance I provided on communicating with schools based on the underlying values of true partnership and respect is paying off. Feedback from the schools on the team is amazing and schools often seem surprised to be treated with respect by researchers and the team. (Research diary 28/02/2007)

I soon realised that I had to deal with this fear and that the solution was not to let go of all responsibility, but to entrench in the team the spirit of our collaboration with the schools and the importance of the contribution of the schools. To allow for further expansion of the project, good training for the entire team is required. Training should not just concentrate on logistics and procedures, but also on the quality of the relationship being sought. Procedures for checking and revising written communications and record keeping of interactions with schools so duplication does not take place are also essential. If the core values of trust and respect are not built into training for the team of a feedback system, these values will be lost during expansion of the system and introduction of new team members.

9.6 Conclusions

Six major conclusions were generated from the research results of this thesis:

1. *An effective feedback system facilitates appropriate use through a gradual process of enlightenment, is flexible and responsive to user inputs, values collaboration and includes instrument, reporting and support components in its design.*

This research shows that appropriate use of feedback is not facilitated by a single event or component in the feedback system, but through the interaction of the various components of the system and repeated exposure of the users to feedback (through for instance reports, feedback sessions and electronic resources). This results in enlightenment, which may lead to further use (Chapter 3). To succeed, a feedback system must be flexible and responsive to user inputs as users are the only true experts on their own context and the context is always in flux. By truly valuing collaboration, feedback becomes a bi-directional process that encourages ownership, trust and respect. This, in turn influences the feedback and monitoring culture to be more receptive and responsive to feedback.

An effective feedback system must also incorporate three components: instruments, reporting and support. The quality of the feedback system is not independent from the monitoring instruments, as appropriate use cannot take place unless the instrument generates reliable data from which valid inferences can be made for the context. Involving users in the development or contextualisation of such instruments helps to establish the credibility of the instruments and the data they generate as discussed in Chapter 3 and supported by authors such as Nevo (2001) and Patton (Patton, 1997). Instrument quality was addressed under sub-question 2 in Chapter 2.

The reporting components are essential to the feedback system and multiple guidelines for effective reporting have been generated in this thesis (Chapters 5-8). The reporting component includes aspects such as feedback sessions, reports and manuals. The reporting must however be supplemented by a support component to facilitate the appropriate use of the feedback. Support

must be provided on two fronts: firstly on the statistical-literacy front in order to help users to understand the data (Earl & Katz, 2006; Schield, 2004) and secondly on the front of data-use (Sub-questions 3-4 in Chapters 5-8).

2. *An optimum feedback system positively influences school data-use culture.*

The school culture of feedback and monitoring use is a major determinant in how receptive and responsive schools are to feedback (Salpeter, 2004). Some authors have argued that this culture need not be seen as a fixed variable in the feedback use equation (Schildkamp & Kuiper, 2009). The monitoring and feedback culture of a school not only relates to previous experiences of data-use, monitoring and feedback, but also to the exposure and level of data-literacy of users as some data-literacy, including statistical literacy is essential in using feedback effectively (Cradler, 2008; Earl & Katz, 2006; Schield, 2004).

This research has shown that a feedback system can positively mediate the feedback and monitoring culture of a school on both these fronts, thereby improving the receptiveness and responsiveness of the school to further feedback. This is firstly accomplished by providing a non-threatening, collaborative atmosphere that encourages respectful communication and positive experiences of feedback use. This approach to feedback is also supported by Bosker, et al.(2007), Brinko (1993) and Hulpia and Valcke (2004). Secondly, the feedback system can help increase user data-literacy by providing examples of interpretation, multiple presentations of data and guidance. This may improve users' sense of self-efficacy and knowledge of data-literacy, thereby encouraging use of feedback and preparing users for future feedback. This conclusion relates to research sub-question 3, Chapters 6-8.

9.3.5 Sub-question 5

How effective is the feedback system in enhancing classroom practices, management and planning activities?

The question relates not only to the expected efficacy of the system, but also to examples of actual efficacy. Cycle 3 of the Prototyping Phase as well as the Assessment Phase related to this question. Actual efficacy was examined in a limited fashion through the observations, interviews and journals kept by the three cases of school use discussed above (see Chapter 7 for a full discussion). During the Assessment Phase, specific questions in the final evaluation questionnaires (management n=13, teachers n=14), as well as the final evaluation reports of the expert evaluators contributed further data on efficacy. The final evaluation questionnaire ratings for both management and teachers indicated that all elements of the feedback system were rated as highly effective with mode ratings of 4 or 5 out of a possible 5 for all elements. The evaluations showed that the data were employed for ***practise*** in the schools, including for instructional purposes, encouraging self directed learning and motivation. The data also contributed to teachers' understanding of their own practice and professional development needs. The data were used by ***management*** for policy development, professional development as well as monitoring of assessment and educational standards. The data were employed in ***planning*** about learners, curriculum planning, professional development planning and for supporting conversations with other stakeholders. The final research question is dealt with in the following section and relates to the design guidelines for the development or adaptation of an effective feedback system.

9.3.6 Sub-question 6

Which design guidelines for the development of an effective feedback intervention for school-based monitoring can be identified?

Design guidelines were developed throughout all three phases of the design research process, based on the various evaluations. The design guidelines

3. *An effective feedback system must offer a comprehensive package to accommodate different users, with various levels of data sophistication, functioning in diverse contexts.*

The data from sub-question 4 (Chapter 7) identified three different approaches to appropriate data-use (Team, Cascade and Top-down approaches) and there may be many more possible approaches. The level of sophistication of the data-use in the schools in this study also varied drastically. It would seem that the most appropriate and effective approach to use may well depend on the culture of the school, school leadership approach, level of teacher development, context and current level of functioning of the school. A more advanced, sophisticated approach to data-use may not always lead to better data-use and may be disheartening and inappropriate in certain contexts (Chapter 7).

Some authors such as Fullan (2006) believe that context has such a large impact that there cannot be one intervention that suits all contexts. It would however be impossible to design and implement different feedback systems for each approach and level of sophistication of data-use. This researcher proposes that an effective feedback system should provide data in a variety of manners (e.g. textual, graphic, tabulated and electronic), with various presentations (e.g. face-to-face, electronic and written) and levels of sophistication, so that schools can use data as per their needs and level of data-literacy.

The feedback system in this thesis thus represents a comprehensive package with data presented in a variety of formats and disseminated through different modes. Users will therefore typically use a core section of the feedback i.e. the summarised school results as well as tabulated learner results and then select which of the other data is appropriate to address the issues with which they are grappling in their own schools. Users may also focus on their preferred mode of feedback to suit their preferences. In this way, one feedback system can be appropriate for multiple user profiles.

4. *An effective feedback system mediates thinking about educational instruction and curriculum.*

All feedback systems have an underlying paradigm in which philosophy, values and pedagogy are made apparent in its various components (Chapters 5-7) The paradigm is present, even if not declared to users. In this research the feedback system aimed at facilitating use. The paradigm therefore encouraged:

- Use of data for evidence-based practise
- The need for differential teaching
- Assessment for learning as opposed to assessment of learning
- Greater understanding of the curriculum

If the paradigm is embodied in each component of the feedback system (from the language used to the congruence between the various elements in the feedback system), the feedback system itself can become a powerful vehicle to bring about changes in how schools think about educational instruction and curriculum. In this instance, the paradigm shift facilitated by the feedback system was a move from merely generating data to interpreting data, planning according to the data, setting goals and evaluating the success of actions. The feedback system therefore acted as a model of the type of appropriate use the system wished to facilitate in schools.

5. *Clear, simple, intuitive data presentation allows for experiential learning to increase user data-literacy.*

Feedback systems often present data in a traditional, academic fashion, which makes high demands of data-literacy, with support aimed at teaching statistical knowledge (Vanhoof, Verhaeghe, Van Petegem, & Valcke, 2010). While some statistical literacy is important (Cradler, 2008; Earl & Katz, 2006), this type of support can be confusing to users and diverts attention from using the data to trying to contend with how the data are presented. This research showed that providing data in a format that is more easily accessible to users through multiple data representations (e.g. graphs, tables and textual discussions) data becomes more accessible and user preferences are

accommodated (Chapters 7-8). This allows the user to focus on use as opposed to trying to understand the data and provides the opportunity to expand data-literacy experientially by interacting with the data. Employing this approach means that users' experiences of data-use must be taken into consideration (Coe, 2002; Hulpia & Valcke, 2004). It is therefore necessary to examine how users already interact with the data (Chapter 7-8). Other systems such as asTTle, where reporting formats were one of the first aspects examined, also support this approach to data representation (Meagher-Lundberg, 2001)

6. *Design research offers an appropriate and powerful approach to adapting, developing and optimising a feedback system.*

Finally, throughout this research the importance of context, consultation and true collaboration with users has been emphasised. Design research is an effective method of developing various prototypes in collaboration with users to ensure contextual appropriateness (Plomp, 2009). The process of participation in the design research itself encouraged ownership and familiarity with the feedback system. Employing a design research approach to develop or adapt a learner performance feedback system to a specific context therefore can have the benefit of improving the receptiveness and responsiveness of schools to feedback.

Design research also provides the opportunity for process use and learning on evaluation and data-use, gained through user participation in evaluations. Design research firmly supports true collaboration between researchers and users. This relationship supports and strengthens respectful, two-way communication (Lachat & Smith, 2005) which should be encouraged in an effective feedback system. The process also generated design principles (to address sub-question 6), which can be employed by other researchers wishing to develop or adapt an effective feedback system in other contexts (see Chapters 5-8)

9.7 Recommendations

The recommendations for this research are divided under research policy and practise. These are now discussed separately.

Research

The design research process for this thesis was ceased after four prototypes. Additional cycles may be required, but the evidence suggests that the feedback system is functioning well in the context for the languages for which it was adapted. Although the design research process was effective in designing the feedback system, it may be necessary to employ a randomised control study to establish fully the effectiveness of the feedback system as suggested by Nieveen (2009) and McKenney, Nieveen and Van den Akker, (2006). Such a study may provide data to support up-scaling of the project.

The feedback system studied here was of limited scale and concentrated on one geographic region, three language groups and only Grade 1 learners. Further research is required to ensure that the principles are transferable to other contexts and grades, language groups and secondary school level. Research that produces design principles on how to roll-out the feedback system in such a way that it remains sustainable and functions well on a large scale is also required.

The feedback and monitoring processes studied in this thesis were externally facilitated. Staff of the CEA administered the SAMP assessments, the data were analysed at the CEA and the CEA provided feedback to the schools. The long-term goal for this feedback system is that it be administered by schools for greater autonomy. This is an important step in making the feedback system part of the formal education system. The characteristics and design of a feedback system, which is internally facilitated and part of the formal education system, may differ from those discussed in this thesis and a further examination of use of feedback in such as system would be beneficial.

This study also only examined three different approaches to effective data-use in schools. A larger study of processes of data-use in schools may

uncover different approaches and identify more elements of effective use. This would provide further information for policy makers on how to support use of data in schools. A study of approaches that result in poor use, misuse and non-use may also provide additional data on how to facilitate use. As not all poor use, misuse and non-use are intentional, such research may provide information on barriers to appropriate use of feedback that exist and how to address them.

Policy

The Whole School Evaluation Policy in RSA already requires schools to participate in self-evaluation and report it to the DoE. Many schools are however, ill-equipped to conduct self-evaluations. The policy should be supported through access to approved existing feedback, monitoring and evaluation systems, which provide the opportunity for schools to develop their own self-evaluative skills.

From this study it is clear that there is not only one approach to effective data-use which would be appropriate for all school cultures, school leadership approaches, levels of teacher development, contexts and various levels of functioning of schools (Chapter 7). It would therefore be counter-productive for policy to prescribe a specific detailed approach to use of feedback and data. The policy should however make provision for time and resources for data-use to take place. The policy on school leadership training should also include training in data-literacy and should guide schools in effective data-use from understanding to evidence-based practise, possibly through exemplars of different approaches for effective data-use.

The current policy (currently under review) on assessment in schools dictates that a large amount of assessment takes place with accountability demands that require teachers spend a significant portions of their time on administration. A shift in policy is recommended, where the emphasis is not on over-assessing and reporting learner data. Fewer assessment opportunities aimed at guiding further teaching and planning are recommended. The decrease in administrative load would provide additional

time to interrogate and interact with the data and provide more time on task in the classroom.

In this thesis it has been shown that ICT has been used effectively internationally to facilitate use of feedback, decrease reporting turn around time and increase access to resources to facilitate use of feedback. In South Africa use of ICT is still limited, not only by challenges of access to ICT infrastructure, but also by limited ICT knowledge and exposure for school staff. It is recommended that the training policy for teachers and staff in schools include training in use of computers and basic data analysis software such as *Excel* as well as on how to access and evaluate resources available on the internet. Development of the ICT infrastructure need not focus on individual schools, but may employ a cluster approach where resources and capacity are improved for a group of schools to share. Such a policy would increase the ability of schools to understand, take ownership and greatly improve access to resources for responding to feedback.

Practice

The recommendations for practice are captured in the section on design guidelines in this chapter (see Section 9.3.6). It is clear that when designing or adapting a feedback system, the level of data-literacy of the users must be considered. Reporting must be planned and piloted with users well in advanced to ensure that data are presented in as clear and intuitive a manner as possible so that users can make sense of the data. All communication should be bi-directional communication between users and facilitators. This type of communication engenders trust, credibility and respect. Data must be represented in several manners and be detailed and have diagnostic value so it may suit the needs of various users. Providing links to resources to support use of feedback is essential to facilitate action. This provides a bridge between understanding and using the data. The culture of the school must also be considered. If ignored, an un-receptive culture may lead to failure of the feedback system, irrespective of its technical quality.

From this research, it is clear that employing evidence-based practice and data-literacy are still a challenge for most teachers in South Africa. It is suggested that training in evidence-based practices as well as basic data-literacy (being able to select and evaluate appropriate data for use, interpreting the data, planning and goal setting as well as evaluating the success of actions based on the data) be introduced into both pre-service and in-service training of teachers.

References

Alkin, M. C., & Daillak, R. H. (1979). A study of evaluation utilization. *Educational Evaluation and Policy Analysis*, *1* (4), 41-49.

Alkin, M. C., & Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation, 29* (1), 1-12.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). New Jersey: Prentice Hall.

Angelle, P. (2004). Monitoring performance through feedback: book review. *School Effectiveness and School Improvement*, *15* (1), 115-120.

Appalsamy, P. (2004). The process and phases of Arthur Inman's suicidal career. *Psychological Society of South Africa*, *34* (2), 237-258.

Archer, E. (2006a). *PIPSSA planning document January 2006.* Pretoria: Centre for Evaluation and Assessment.

Archer, E. (2006b). *Proposed changes to PIPS paper-based baseline assessment for the South African context.* Pretoria: Unpublished.

Archer, E., Scherman, V., Coe, R., & Howie, S. (2010). Finding the right fit: the adaptation and translation of the Performance Indicators for Primary Schools for the South African context. *Perspectives in Education*, *28* (1), 77-88.

Atlantech. (2003). *asTTle assists teaching and learning across New Zealand*. Retrieved 2008 25, February, from www.atlantech.co.nz/AstlleassiststeachingandlearningacrossNewZealand_489d47d0d0.html

Babbie, E., & Mouton, J. (2001). *The practice of social research: South African edition.* Oxford: Oxford University Press.

Baker, F. B. (2001). *The basics of item response theory (2nd ed)*. Retrieved November 15, 2004, from http://ericae.net

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13* (1).

Barry, D. M. (2005). *PIPSSA on entry baseline assessment: Report.* Pretoria: Unpublished.

Best, J. W., & Kahn, J. V. (2006). *Research in education* (10th ed.). Boston: Pearson Education, Inc.

Beyer, J. M. (1997). Research utilization: Bridging the gap between communities. *Journal of Management Inquiry*, *6*, 17-22.

Black, P., & Wiliam, D. (1998). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, *80* (2), 139–148.

Bosker, R. J., Branderhorst, E. M., & Visscher, A. (2007). Improving the utilisation of management information systems in secondary schools. *School Effectiveness and School Improvement*, *18* (4), 451 – 467.

Brinko, K. T. (1993). The practice of giving feedback to improve teaching: What is effective? *The Journal of Higher Education*, *64* (5), 574-593.

Brown, G. T. (2007, 19 February). Conceptions of assessment. Lecture presented at Centre for Evaluation and Assessment, University of Pretoria, Pretoria.

Brown, G. T. (2008). Assessment literacy training and teachers' conceptions of assessment. In C. Rubie-Davies, & C. Rawlinson (Eds.), *Challenging Thinking about Teaching and Learning* (pp. 269-285). New York: Nova Science.

Brown, G. T., & Rodger, S. (1999). Research utilization models: Frameworks from implementing evidence-based occupational practice. *Occupational Therapy International*, *6* (1), 1-23.

Brown, G. T., Irving, S. E., & Keegan, P. J. (2008). *An introduction to educational assessment, measurement and evaluation: improving the quality of teacher-based assessment* (2nd ed.). Auckland: Pearson Education New Zealand, Ltd.

Cassassus, J. (2001). School evaluation in Chile. *Prospects*, *31* (4), 565-574.

Chamberlain, M. (2007). *Reading literacy in New Zealand: An overview of New Zealand's results from the Progress in International Reading Literacy Study (PIRLS) 2005/2006.* Wellington: Ministry of Education.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis.* London: Sage Publications Ltd.

Christie, C. A., & Alkin, M. C. (1999). Further reflections on evaluation misutilization. *Studies in Educational Evaluation*, *25*, 1-10.

Clandinin, D. J., & Connelly, F. M. (1994). Personal experience methods. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 413-427). Thousand Oaks: Sage Publications Ltd.

Coe, R. (2002). Evidence on the role and impact of performance feedback in schools. In A. J. Visscher, & R. Coe (Eds.), *School improvement through performance feedback* (pp. 3-26). Lisse: Swets & Zeitlinger Publishers.

Coe, R., & Visscher, A. J. (2002). Drawing up the balance sheet for school performance feedback systems. In A. J. Visscher, & R. Coe (Eds.), *School improvement through performance feedback* (pp. 221-254). Lisse: Swets & Zeitlinger Publishers.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). New York: Routledge Falmer.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design Research: Theoretical and Methodological Issues. *The Journal Of The Learning Sciences*, *13* (1), 15-42.

Consortium for Reseach on Education, Access, Transitions & Equity (CREATE). (2009, August). *'No Fee' schools in South Africa: policy brief number 7.* Retrieved June 9, 2010, from Create-rpc: http://www.create-rpc.org/PDF_documents/Policy_Brief_7.PDF

Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, *56* (3), 331-364.

Cradler, J. (2008). *Data-driven decision-making and the electronic learning assessment resources (ELAR).* Retrieved May 12, 2010, from http://www.clrn.org/elar/dddm.cfm

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American psychologist*, *30*, 116-127.

Crooks, T. J. (2002). Educational assessment in New Zealand schools. *Assessment in Education*, *9* (2), 237-253.

Curriculum Evaluation and Management Centre. (2002). *Reliabilities*. Retrieved December 3, 2002, from http://www.midyisproject.org/reliabilities.asp

Curriculum Evaluation and Manangement Centre. (2007). Retrieved April 10, 2007, from CEM home page: http://www.cemcentre.org

Czarniawska, B. (2004). *Narratives in social science research.* London: Sage Publications.

Data Management and Analysis Division, New Zealand Ministry of Education. (2008, December). *Education Statistics of New Zealand 2007*. Retrieved July 15, 2010, from Education Counts: http://www.educationcounts.govt.nz/publications/ece/2507/27456/3

De Villiers, M. R. (2005). Three approaches as pillars for interpretive information systems research: development research, action research and grounded theory. *Research for a changing world: proceedings of South African Institute for Computer Scientists and Information Technologists 2005 conference* (pp. 111-120). White River: SAICSIT.

Department of Education. (2001). *Report on the school register of needs 2000 survey.* Pretoria: Department of Education.

Department of Education. (2002a). *Education statistics in South Africa at a glance 2000.* Pretoria: EMIS directorate, Department of Education.

Department of Education. (2002b). *Grade 3 National systemic evaluation report.* Pretoria: Government printers.

Department of Education. (2002c). *Revised National Curriculum Statement Grades R-9 (schools)*. Retrieved October 25, 2006, from http://www.polity.org.za/PDF/Curriculum.PDF

Department of Education. (2003a). *Develop an HIV & AIDS plan for your school: A guide for school governing bodies and management teams.* Pretoria: Department of Education.

Department of Education. (2003b). *Draft white paper on e-education: transforming learning and teaching through information and communication technologies.* Pretoria: Government Printers.

Department of Education. (2003c). *Education statistics in South Africa at a glance 2001.* Pretoria: EMIS directorate, Department of Education.

Department of Education. (2004). *Education statistics in South Africa at a glance 2002.* Pretoria: Department of Education.

Department of Education. (2005a). *Education statistics in South Africa at a glance 2003.* Pretoria: Department of Education.

Department of Education. (2005b). *Education statistics in South Africa at a glance 2004.* Pretoria: Department of Education .

Department of Education. (2006a). *Grade 3 Foundation Phase National systemic evaluation report.* Pretoria: Government Printer.

Department of Education. (2006b). *Grade 6 Intermediate Phase Systemic Evaluation Report.* Pretoria: Government Printer.

Department of Education. (2006c, November 22). Provinces submit their list of no-fee schools for 2007. [Media statement].

Department of Education. (2006d). *The national policy framework for teacher education and development in South Africa: more teachers; better teachers.* Pretoria: Department of Education.

Department of Education. (2007a). *2007 School Realities .* Pretoria: Government Printers.

Department of Education. (2007b). *National assessment report (public ordinary schools): National Education Infrastructure Management System (NEMIS).* Pretoria: Department of Education.

Department of Education. (2009). *NEMIS (National Education Infrastructure Management System) PDF Report 2009.* Pretoria: Department of Education.

Department of Education. (n.d.). *National Curriculum Statement General Education and Training, Assessment guidelines: assessment guidelines for Foundation phase grades r - 3.* Pretoria: Government Press.

Dunham, B. R. (1995). *The DELPHI technique*. Retrieved May 5, 2008, from http://www.medsch.wisc.edu/adminmed/2002/orgbehav/delphi.PDF

Earl, L. M., & Katz, S. (2006). *Leading schools in a data-rich world.* Thousand Oaks: Corwin Press.

Education Labour Relations Council. (2003). *Training manual on the Integrated Quality Management System*. Retrieved April 16, 2007, from

www.elrc.co.za/UploadedDocuments/IQMS%20Training%20Manual.doc

Einstein, A. (1954). *Ideas and Opinions.* New York: Crown Publishers, Inc.

Eiselen, S. (2004). *PIPSSA Handover document.* Pretoria: Unpublished.

Eiselen, S. (2005a). Items included for analysis for November 2005 (PIPSSA). Pretoria: [Unpublished memo].

Eiselen, S. (2005b). *PIPSSA Future plans 2005.* Pretoria: Unpublished.

Estabrooks, C. A. (1999). The conceptual structure of research utilization. *Research in Nursing & Health*, *22*, 203–216.

Fiske, E. B., & Ladd, H. F. (2001). Self-governing schools and accountability in New Zealand. *Prospects*, *31* (4), 537-552.

Fitz-Gibbon, C. T., & Tymms, P. (2002). *Technical and ethical issues in indicator systems: Doing things right and doing wrong things.* Retrieved December 15, 2009, from Education Policy Analysis Archives10 (no. 6): http://epaa.asu.edu/epaa/v10n6/

Fitz-Gibbon, C. T. (1996). *Monitoring education: indicators, quality and effectiveness.* London: Continuum.

Frisbee, D. A. (1988). Reliability of scores from teacher-made tests. *Educational Measurement: Issues and Practice*, *7* (1), 25-35.

Fuchs, L. S., Fuchs, S., Eaton, S. B., Hamlett, C. L., & Karns, K. M. (2000). Supplementing teacher judgements of mathematics test accommodations with objective data sources. *School Psychology Review*, *29* (1), 65-85.

Fullan, M. (2006). *Turnaround leadership.* San Francisco: Josie-Bass.

Gardner, J. (2006). *Assessment and learning.* London: Sage Publications.

Gauteng Department of Education. (2005). *Gauteng online to reach 1000 Gauteng Public Schools*. Retrieved January 18, 2006, from http://www.education.gpg.gov.za/publications/gauteng%20online%20to%20reach%20public%20schools.htm

Gauteng Department of Education. (n.d.). *Improving the quality of education in Gauteng*. Retrieved August 7, 2007, from http://www.gautengonline.gov.za/miscimages/003.Ofsted.PDF

Geertz, C. (1975). *The Interpretation of Cultures.* New York: Basic Books.

Gergen, K. J., Lightfoot, C., & Sydow, L. (2004). Social construction: Vistas in clinical child and adolescent psychology. *Journal of clinical child and adolescent psychology*, *33* (2), 389-399.

Glasswell, K., Parr, J., & Aikman, M. (2001). *Development of the asTTle writing assessment rubrics for scoring extended writing tasks. Project asTTle Technical Report #6.* University of Auckland: Ministry of Education.

Goodson, I., & Sikes, P. (2001). *Doing qualitative research in educational settings: learning from lives.* Buckingham: Open University Press.

Greene, J. C. (2007). *Mixed methods in social inquiry: mixing method.* Hoboken: John Wiley & Sons, Ltd.

Gronlund, N. E. (1998). *Assessment of student achievement* (6th ed.). Boston: Allyn & Bacon.

Guba, E. G., & Lincoln, Y. S. (1985). *Understanding and doing naturalistic inquiry.* Beverly Hills: Sage Publications.

Harmon, M., Smith, T. A., Martin, M. O., Kelly, D. L., Beaton, A. E., Mullis, I. V., et al. (1997). *Performance assessment in IEA's Third International Mathematics and Science Study (TIMSS).* Boston College: TIMSS International Study Centre.

Harnar, M. A., & Preskill, H. (2007). Evaluator's descriptions of process use: An exploratory study. *New Directions for Evaluation*, *116*, 27 – 44.

Harris, L., Chapman, C., Muijs, D., Russ, J., & Stoll, L. (2006). Improving schools in challenging circumstance: Exploring the possible. *School Effectiveness and School Improvement*, *17* (4), 399-407.

Hattie, J. (2005, August 7-9). *What is the nature of evidence that makes a difference to learning* . Paper presented at the 2005 Acer Research Conference, Melbourne, Australia Retrieved March 7, 2006, from www.acer.edu.au/workshops/documents/Hattie.PDF.

Hattie, J. (2009, April). Visibly learning from reports: The validity of score reports. *Paper presented at the annual meeting of the National Council on Measurement in Education (NCME)* . San Diego, CA.

Hattie, J. A., & Brown, G. T. (2004). *Cognitive processes in asTTle: The SOLO taxonomy. asTTle Technical Report #43.* University of Auckland: Ministry of Education.

Hattie, J. A., & Brown, G. T. (2008). Technology for school-based assessment and assessment for learning: Development principles from New Zealand. *Journal of Educational Technology Systems*, *26* (2), 189-201.

Hattie, J. A., Brown, G. T., & Keegan, P. J. (2003). A national teacher-managed, curriculum-based assessment system: Assessment Tools for Teaching & Learning asTTle). *International Journal of Learning*, *10*, 771-778.

Hattie, J. A., Brown, G. T., Keegan, P. J., MacKay, A. J., Irving, S. E., Cutforth, S., et al. (2004a). *Assessment Tools for Teaching and Learning (asTTle) Manual (Version 4, 2005).* Wellington, NZ: University of Auckland/ Ministry of Education/ Learning Media.

Hattie, J. A., Brown, G. T., Keegan, P. J., MacKay, A. J., Irving, S. E., Patel, P., et al. (2004b). *Assessment Tools for Teaching and Learning (asTTle) Version 4, 2005: Manual.* Wellington, New Zealand: University of Auckland: Ministry of Education Learning Media.

Hattie, J. A., Brown, G. T., Ward, L., Irving, S. E., & Keegan, P. J. (2006). Formative evaluation of an educational assessment technology innovation: Developers' insights into Assessment Tools for Teaching and Learning (asTTle). *Journal of Multi-Disciplinary Evaluation*, *5* (3), 1-54.

Hattie, K., & Jaeger, R. (1998). Assessment and classroom learning: a deductive approach. *Assessment in Education: Principles, Policy & Practice*, *5* (1), 111-123.

Hayward, L., & Hedge, N. (2005). Travelling towards change in assessment: policy, practice and research in education. *Assessment in Education*, *12* (1), 55-75.

Hendriks, M. A., Doolard, S., & Bosker, R. J. (2001). School self-evaluation in the Netherlands: development of the ZEBO-instrumentation. *Prospects*, *31* (4), 503-518.

Henning, E., Van Rensburg, W., & Smit, B. (2004). *Finding your way in qualitative research.* Pretoria: Van Schaik.

Hood, J. (1990). *Education: Is America spending too much?* Retrieved August 7, 2007, from Cato policy analysis, 126: http://www.cato.org/pubs/pas/pa126.html

Howie, S. J. (1997). *Mathematics and Science Performance in the Middle School Years in South Africa.* Pretoria: Human Sciences Research Council.

Howie, S. J. (2001). *Mathematics and Science Performance in Grade 8 in South Africa 1998/1999. TIMSS-R in South Africa.* Pretoria: Human Sciences Research Council.

Howie, S. J. (2002). *English language proficiency and contextual factors influencing mathematics achievement of secondary school pupils in South Africa.* Enschede: University of Twente.

Howie, S. J., & Plomp, T. (2005). International comparative studies of education and large scale change. In N. Bascia, A. Cummings, K. Datnow, K. Leithwood, & D. Livingstone (Eds.), *International handbook of educational policy* (pp. 75-100). Oxford: Springer.

Howie, S. J., Venter, E., Van Staden, S., Zimmerman, L., Long, C., Scherman, V., et al. (2008). *Progress in International Reading Literacy Study (PIRLS) 2006 summary report: South African children's reading literacy achievement.* Pretoria: University of Pretoria.

Hulpia, H., & Valcke, M. (2004). The Use of Performance Indicators in a School Improvement Policy: The Theoretical and Empirical Context Department of Education, Ghent University, Belgium. *Evaluation and Research in Education*, *1&2*, 102-119.

Human Science Research Council. (2006). *Mathematics and science achievement in South Africa, TIMSS 2003.* Pretoria: HSRC press .

Husėn, T., & Tuijnman, A. C. (1994). Monitoring standards in education: why and how it came about. In A. C. Tuijnman, & T. N. Postlethwaite (Eds.), *Monitoring the standards of education* (pp. 1-22). Oxford: Pergamon .

Illinois institute of technology. (nd). *The DELPHI method: Definition and Historical Background*. Retrieved May 5, 2008, from http://www.iit.edu/~it/delphi.html

Jansen, J. D. (2001). On the politics of performance in South African education: Autonomy, accountability and assessment. *Prospects*, *31* (4), 553-564.

Johnson, R. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning*, *21*, 93-110.

Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). Thousand Oaks, CA: Sage publications, Inc.

Joseph, D. (2004). The Practice of Design-Based Research: Uncovering the Interplay Between Design, Research, and the Real-World Context. *Educational Psychologist*, *39* (4), 235 - 242.

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17-64). Westport, CT: American Council on Education and Praeger.

Kelly, A. E. (2003). Research as Design. *Educational Researcher*, *32* (1), 3-4.

Killen, R. (2003). Validity in outcomes-based education. *Perspectives in Education*, *21* (1), 1-14.

King, J. A., & Pechman, E. M. (1984). Pinning a wave to the shore: Conceptualizing evaluation use in school systems. *Educational Evaluation and Policy Analysis*, *6* (3), 241-251.

Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation*, *88*, 5-23.

Kline, P. (1993). *The handbook of psychological testing.* London: Routledge.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback onterventions on performance: A historical review, a ,eta-Analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119* (2), 254-284.

Krathwohl, D. R. (1998). *Methods of educational and social science research: An integrated approach* (2nd ed.). New York, Longman.

Küstner, P. M., & Eiselen, S. (2005). *Notes on PIPSSA.* Pretoria: Unpublished.

Lachat, M. A., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed at Risk*, *10* (3), 333-349.

Lewins, A., & Silver, C. (2009, April 2). *Choosing a CAQDAS Package.* Retrieved May 5, 2009, from CAQDAS Networking Project and Qualitative Innovations in CAQDAS Project (QUIC): www.caqdas.soc.surrey.ac.uk.htm

Linn, R. L., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). New Jersey: Prentice Hall.

Looney, A. (2006). Assessment in the Republic of Ireland. *Assessment in Education*, *13* (3), 345 - 353.

Love, J. M. (1985). Knowledge transfer and utilization in education. *Review of Research in Education*, *12*, 377-386.

Maddux, J. E., Snyder, C. R., & Lopez, S. J. (2004). Towards a positive clinical psychology: Deconstructing the illness ideology and constructing ideology of human strengths and potential. In P. A. Linley, & S. Joseph (Eds.), *Positive psychology in practice* (pp. 320-353). New Jersey: John Wiley & Sons, inc.

Mafumiko, F. S. (2006). *Micro-scale experimentation as a catalyst for improving the chemistry curriculum in Tanzania.* Enschede: University of Twente.

Martin, M. O., Mullis, I. V., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 International Science Report: Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades.* Boston College: TIMSS & PIRLS International Study Center.

Martin, M. O., Mullis, I. V., Gonzalez, E. J., Gregory, K. D., Smith, T. A., Chrostowski, S. J., et al. (2000). *TIMSS 1999 International Science Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade.* Boston College: The International Study Centre.

McKenney, S. (2001). *Computer-based support for science education materials developers in Africa: exploring potentials* . Enschede: University of Twente.

McKenney, S., Nieveen, N., & Van den Akker, J. (2006). Educational design research. In J. Van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Design research from a curriculum perspective* (pp. 62-90.). London: Routledge.

Meagher-Lundberg, P. (2001). *Output reporting design: Focus group 1. (Technical Report 9).* Auckland: University of Auckland: Project asTTle.

Merrell, C., & Tymms, P. (2005). International comparisons of early literacy. *Literacy today*, *2005* (March), 18-19.

Mertens, D. M. (1998). *Research methods in education and psychology: integrating diversity with quantitative and qualitative approaches.* Thousand Oaks, CA: SAGE Publications.

Mertens, D. M. (2009). *Transformative research and evaluation.* New York: Guilford Press.

Messick, S. (1989). Meaning and values in test validation: the science and ethics of assessment. *Educational Researcher*, *18* (2), 5-11.

Michigan State University Extension. (1994). *DELPHI technique*. Retrieved May 5, 2008, from http://web1.msue.msu.edu/msue/imp/modii/iii00006.html

Muller, J. (2005). The challenge of cognitive demand. In Matric: What is to be done? *Seminar papers and presentations 23 June 2005* (pp. 40-44). Johannesburg: Umalusi.

Mullis, I. V., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 International Mathematics Report: Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades.* Boston College: TIMSS & PIRLS International Study Center.

Mullis, I. V., Martin, M. O., Gonzalez, E. J., Gregory, K. D., Garden, R. A., O'Connor, K. M., et al. (2000). *TIMSS 1999 International Mathematics Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade.* Boston College: International Study Centre.

Murphy, K. R., & Davidshofer, C. O. (1994). *Psychological testing: principles and applications.* New Jersey: Prentice-Hall, Inc.

National Treasury Republic of South Africa. (2005). Retrieved April 21, 2006, from Provincial budgets and Expenditure review: 2001/02-2007/2008 September 2005: www.treasury.gov.za

Nevo, D. (2001). School evaluation: internal or external? *Studies in Educational Evaluation*, *27*, 95-106.

Nevo, D. (2007, 23 March). International comparisons of student achievements: for good and for worse – the case of Israel. *Lecture presented at the Centre for Evaluation and Assessment* . University of Pretoria, Pretoria.

New Zealand Ministry of Education. (1994). *Assessment: Policy to practice.* Wellington: Learning Media, Ltd.

New Zealand Ministry of Education. (2007a). *E-asTTle FAQ*. Retrieved April 10, 2007, from http://www.tki.org.nz/r/asttle/e-asttle/faq_e.php

New Zealand Ministry of Education. (2007b). *PISA (Programme for International Student Assessment)*. Retrieved January 29, 2008, from www.educationcounts.govt.nz/data_collections/pisa_research

New Zealand Ministry of Education. (2009). *home*. Retrieved October 18, 2009, from TKI-asTTle: http://www.tki.org.nz/r/asttle

New Zealand Ministry of Education. (n.d.). *TIMSS (Trends in International Mathematics and Science Study) Home Page*. Retrieved January 29, 2008, from www.educationcounts.govt.nz/data_collections/timss

Nicholls, H. (2003). *English curriculum framework & map: Levels 2—6. Project asTTle Technical Report #34.* University of Auckland: Ministry of Education.

Nieveen, N. (1997). *Computer support for curriculum developers: A study of the potential of computer support in the domain of formative curriculum evaluation.* Enschede: PrintPartners Ipskamp.

Nieveen, N. (2009). Formative evaluation in educational design research. In T. Plomp, & N. Nieveen (Eds.), *An introduction to educational design research* (pp. 89-102). Enschede, Netherlands: SLO - Netherlands institute for curriculum development.

Nuttal, D. L. (1994). Monitoring national standards: United Kingdom. In T. Husén, & T. N. Postlethwaite (Eds.), *The International encyclopaedia of education* (pp. 3906-3912). Oxford: Pergamon.

Onwuegbuzie, A. J. (2002). Why can't we get along? Towards a framework for unifying research paradigms. *Education*, *122* (3), 518-530.

Onwuegbuzie, A. J., & Johnson, R. B. (2004). Mixed methods research: a research paradigm whose time has come. *Educational Researcher, 33* (7), 14-26.

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue of mixed research. *Research in the Schools*, *13* (1), 48-63.

Onwuegbuzie, A. J., & Leech, N. L. (2005). Taking the ''Q'' out of research: teaching research methodology courses without the divide between quantitative and qualitative paradigms. *Quantity & Quality*, *39*, 267-296.

Organisation for Economic Co-Operation and Development (OECD). (2008). *Review of national policies for education South Africa.* Paris: OECD publications.

Owen, K., & Taljaard, J. J. (1996). *Handbook for the use of psychological and scholastic tests of the HSRC.* Pretoria: Human Science Research Council.

Pandor, N. (2006, December 28). Not where we want to be. *Statement by Naledi Pandor, MP, Minister of Education, on the release of the 2006 senior certificate examination results* . Pretoria: Parliament.

Parliamentary Monitoring Group. (2006, June 20). *Report to the portfolio commission on the Integrated Quality Management system*. Retrieved April 16, 2007, from http://www.pmg.org.za/docs/2006/060620iqms.htm

Patton, M. Q. (1991). Towards utility in reviews of multivocal literatures. *Review of Educational Research*, *61* (3), 287-292.

Patton, M. Q. (1997). *Utilization-focused evaluation: the new century text* (3rd ed.). California: SAGE Publications, Inc.

Plomp, T. (2009). Educational Design Research: an Introduction. In T. Plomp, & N. Nieveen (Eds.), *An Introduction to Educational Design Research* (pp. 9-37). Enschede: SLO.

Plomp, T., Howie, S. J., & McGaw, B. (2003). International studies in educational achievement. In T. Kellaghan, & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 951-978). Dordrecht: Kluwer Academic publishers.

Plummer, K. (2001). *Documents of life 2: an invitation to a critical humanism.* London: SAGE Publications.

Powell, A., & Maritz, J. (2006, January 19). Rasool ready to crack whip on bad schools. *The Cape Times*, p. 1.

Reeves, T. C., Herrington, J., & Oliver, R. (2005). Design Research: A Socially Responsible Approach to Instructional Technology Research in Higher Education. *Journal of Computing in Higher Education*, *16* (2), 97-116.

Richey, R. T., Klein, J. D., & Nelson, W. A. (1996). Developmental research. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 1213-1245). New York: Macmillan.

Riessman, C. K. (1993). *Narrative Analysis.* Newbury Park: SAGE Publications.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: a systemic approach* (7th ed.). Thousand Oaks: Sage Publications, Inc.

Rowe, K. J., Turner, R., & Lane, K. (2002). The VCE Data Project. In A. J. Visscher, & R. Coe, *School improvement through performance feedback* (pp. 163-190). Lisse: Swets & Zeitlinger Publishers.

Saldana, J. (2009). *The coding manual for qualitative researchers.* London: SAGE Publications Ltd.

Salpeter, J. (2004). Data: Mining with a mission. *Technology*, *March*, 30-36.

Saunders, L. (2001). The use of 'value-added' measures in school evaluation: a view from England. *Prospects*, *31* (4), 489-502.

Scheerens, J., Glas, C., & Thomas, S. M. (2003). *Educational evaluation, assessment, and monitoring: a systemic approach.* Netherlands: Swets & Zelinger Publishers.

Scherman, V., Archer, E., Howie, S. J., & Lopez, T. (2006). *National Research Foundation: Investigating Value-added assessment in South Africa,*

Technical report year 2003/2004. Pretoria: Centre for Evaluation and Assessment.

Schield, M. (2004). Information literacy, statistical literacy and data literacy. *IASSIST*, *3*, 6-11.

Schildkamp, K. (2007). *The utilisation of a self-evaluation instrument for primary education.* Enschede: Printpartners Ipskamp.

Schildkamp, K., & Kuiper, W. (2009). Data-informed curriculum reform: Which data, what purposes,and promoting and hindering factors. *Teaching and Teacher Education*, *xxx*, 1-15.

Schildkamp, K., & Teddlie, C. (2008). School performance feedback systems in the USA and in the Netherlands: A comparison. *Educational Research and Evaluation*, *14* (3), 255-282.

Schildkamp, K., Visscher, A., & Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *School Effectiveness and School Improvement*, *20* (1), 69–88.

Schildkamp, K., Visscher, A., & Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *School Effectiveness and School Improvement*, *20* (1), 69–88.

Scientific Software Development. (1997). *Atlas.ti the knowledge workbench: Short user's manual.* Berlin: Thomas Muhr, Scientific Software Development.

Scientific Software Development. (2004). *Atlas.ti the knowledge workbench: Quick tour for beginners V5.0.* Berlin: Thomas Muhr, Scientific Software Development.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Chicago, IL: Rand McNally.

Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: theory, research and practise since 1986. *American Journal of Evaluation*, *18* (3), 195-208.

Singh, T. (2004). *School - Based Assessment: the interface between Continuous Assessment (CASS) and the external summative examination at Grade 12 level with special focus on Mathematics*

and Science. Retrieved May 27, 2007, from University of Pretoria: http://upetd.up.ac.za/thesis/available/etd-12062004-141525.

Smith, R. M. (2004). Detecting item bias in the Rasch model. In J. E. Smith, & R. M. Smith (Eds.), *Introduction to Rasch measurement: Theory, models and applications.* Maple Grove: JAM Press.

Statistics South Africa. (2001). *Census 2001: Census in brief.* Pretoria: Statistics South Africa.

Statistics South Africa. (2004). *General household Survey July 2004.* Pretoria: Statistics South Africa.

Statistics South Africa. (2008). *Household survey 2007.* Pretoria: Stats SA.

Suen, H. K. (1990). *Principle of test theories.* New Jersey: Lawrence Erlbaum Associates.

Sukhraj, P. (2006, January 5). Outrage as all matrics at school fail. *The Star*, p. 3.

Syque. (2007). *The Quality Toolbook: How to understand the Nominal Group Technique.* Retrieved April 18, 2008, from http://syque.com/quality_tools/toolbook/NGT/ngt.htm

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches.* Thousand Oaks, CA: SAGE Publications.

Taut, S. M., & Alkin, C. M. (2003). Program staff perceptions of Barriers to evaluation implementation. *American Journal of Evaluation*, *24*, 213-226.

Taylor, N. (2007, June 26-29). Systemic Assessments and Social Justice: Strange Bedfellows, or Not? *Colloquium at the 2007 14th International Learning Conference* . Johannesburg, South Africa.

Taylor, N., Muller, J., & Vinjevold, P. (2003). *Getting schools working: research and systemic school reform in South Africa.* Cape Town: Pearson Education South Africa.

Teddlie, C., Kochan, S., & Taylor, D. (2002). The ABC+ model for school diagnosis, feedback, and improvement. In A. J. Visscher, & R. Coe (Eds.), *School improvement through performance feedback* (pp. 75-114). Lisse: Swets & Zeitlinger Publishers.

Thijs, A. (1999). *Supporting science curriculum reform in Botswana: The potential of peer coaching.* Enschede: PrintPartners Ipskamp.

Tymms, P. (1999). *Baseline assessment and monitoring in primary schools: achievements, attitudes and value-added indicators.* London: David Fulton Publishers.

Tymms, P., & Albone, S. (2002). Performance indicators in primary schools. In A. J. Visscher, & R. Coe (Eds.), *School improvement through performance feedback* (pp. 191-220). Lisse: Swets & Zeitlinger Publishers.

Tymms, P., & Coe, R. (2003). Celebrating the success of distributed research with schools: the CEM Durham. *British Educational Research Journal*, *29* (5), 639-653.

Tymms, P., & Wylde, M. (2003). Baseline assessment and monitoring in primary schools. *Paper presented at the Symposium Connectable Processes in Elementary and Primary Section.* Bamberg - April 2003.

Tymms, P., Merrell, C., & Henderson, B. (2000). Baseline assessment and progress during the first three years at school. *Educational Research and Evaluation*, *6* (2), 105-129.

Tymms, P., Merrell, C., & Jones, P. (2004). Using baseline assessment data to make international comparison. *British Educational Research Journal*, *30* (5), 673-689.

Umalusi. (2007). *The statistical working group*. Retrieved April 16, 2007, from http://www.umalusi.org.za

Umalusi. (2009). *2008 Maintaining standards report.* Pretoria: Umalusi.

United Nations Children's Fund (UNICEF). (2000). *Defining quality in education.* New York: United Nations Children's Fund.

Van den Akker, J. (1999). Principles and Methods of Development Research. In J. J. Van den Akker, R. Branch, K. Gustafson, N. M. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training.* Dordrecht: Kluwer Academic Publishers.

Van den Berg, S., & Burger, R. (2002, October 22-24). Education and socio-economic differentials:A study of school performance in the Western Cape. Paper presented at the 2002 DPRU/FES

conference, Johannesburg, South Africa Retrieved June 9, 2010, from www.commerce.uct.ac.za/Research_Units/DPRU/Conf2003PDF/Conf2002PDF/VdBergBurger.PDF.

Van Petegem, P., & Vanhoof, J. (2005). Feedback of performance indicators: A tool for school improvement? Flemish case studies as a starting point for constructing a model for school feedback. *Revista Eletronca Iberoamericana sobre Calidad, Eficacia y Cambio en EducaciÓn, 3* (1), 222-234, Retrieved January 30, 2007, from http://www.ice.deusto.es/rinace/reice/vol3n1_e/VanPetegemVanhoof.PDF.

Vanhoof, J., Verhaeghe, G., Van Petegem, P., & Valcke, M. (2010). The Influence of Competences and Support on School Performance Feedback Use. *ICSEI 2010.* Kuala Lumpur: ICSEI.

Visscher, A. J. (2002). A framework for studying school performance feedback systems. In A. J. Visscher, & R. Coe (Eds.), *School improvement through performance feedback* (pp. 41-74). Lisse: Swets & Zeitlinger Publishers.

Visscher, A. J., & Coe, R. (2002). Introduction. In A. J. Visscher, & R. Coe (Eds.), *School improvement through performance feedback* (pp. xi-xix). Lisse, Swets & Zeitlinger Publishers.

Voster, C. (2003). *General systems theory and psychotherapy: beyond postmodernism.* Pretoria: SATORI publishers.

Ward, L., Hattie, J. A., & Brown, G. T. (2003). *The evaluation of asTTle in schools: The power of professional development. asTTle Technical Report, #35.* University of Auckland: Ministry of Education.

Weiss, C. H. (1981). Measuring the use of evaluation. In J. A. Ciarlo (Ed.), *Utilizing evaluation: Concepts and measurement techniques.* Beverly Hills: Sage.

Williams, P. L., & Webb, C. (1994). The Delphi technique: a methodological discussion. *Journal of Advanced Nursing*, *19*, 180-186.

Willig, C. (2001). *Introducing qualitative research in psychology: adventures in theory and method.* Philadelphia: Open University Press.

Wohlstetter, P., Datnow, A., & Park, V. (2008). Creating a system for data-driven decision-making: Applying the principal agent framework. *School Effectiveness and School Improvement*, *19* (3), 239-259.

World Bank. (2008). *2008 World Development Indicators.* Washington: World Bank.

World Bank. (2010, June 10). *World Development Indicators* . Retrieved July 15, 2010, from http://www.google.com/publicdata?ds=wb-wdi&met=sp_pop_totl&idim=country:NZL&dl=en&hl=en&q=population+New+Zealand#met=sp_pop_totl&idim=country:NZL:ZAF

Yin, R. K. (2003). *Case study research: design and methods* (3 rd ed.). Thousand Oaks: SAGE Publications.

Zapawa, T. (2005). *Excel Advanced Report Development* (2 nd ed.). Indianapolis: Wiley Publishing, Inc.

Zuzovsky, R. (1994/1995). Utilization of research findings: a matter of tradition. *Knowledge and Policy*, *7* (4), 78-94.