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.
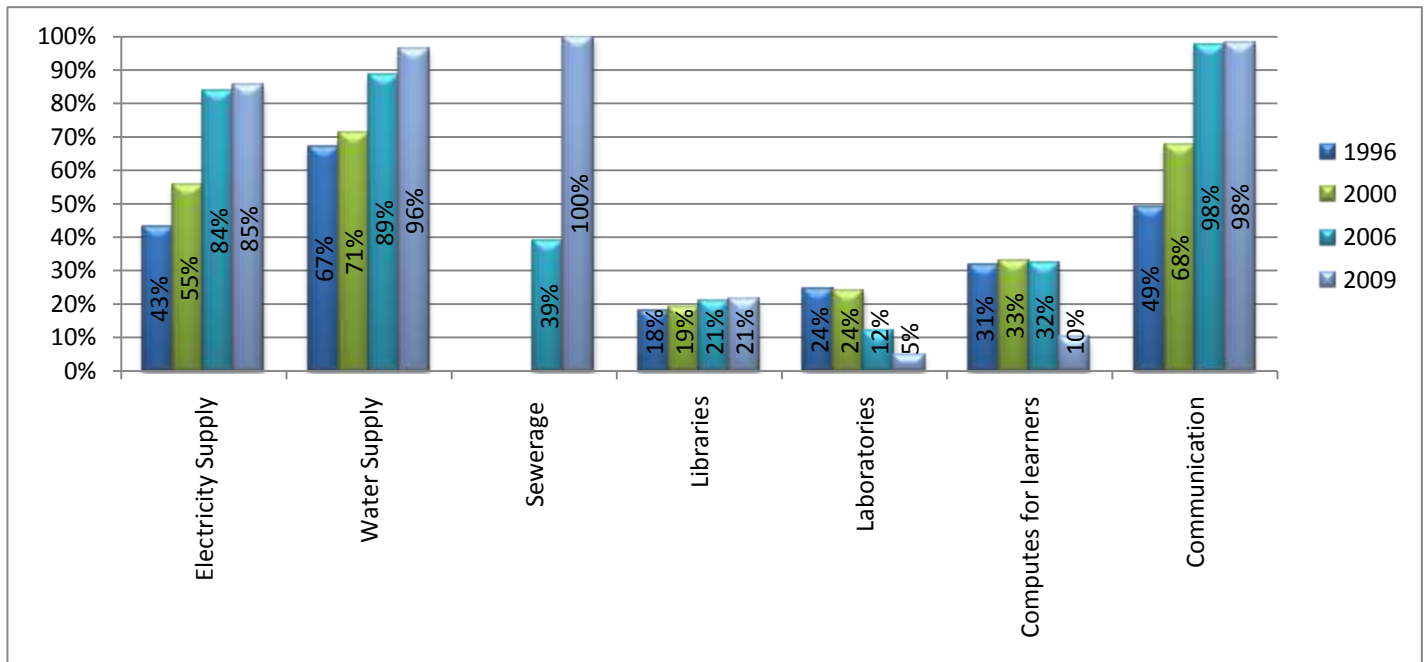

Figure 3.1: Infrastructure changes in the South African education system 1996-2006[17] (public ordinary schools)
(Department of Education, 2001, 2007b, 2009)

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.
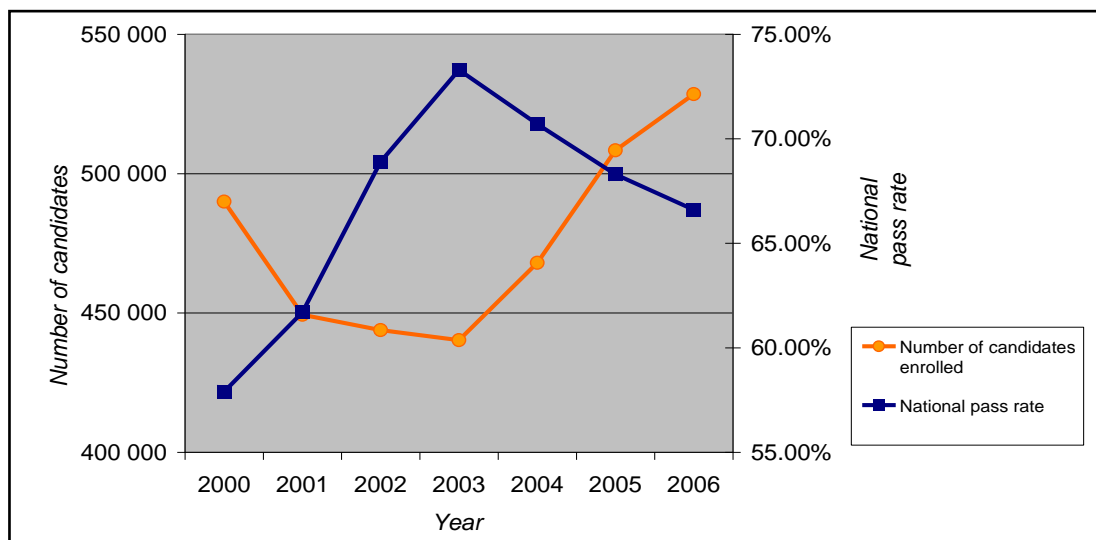

Figure 3.2: Grade 12 pass rate per year mapped against with number of candidates enrolled
Compiled from: Department of Education, 2002a, p. 181, 2003c, p. 222, 2004, p. 183, 2005a, p. 204, 2005b, p. 225; Pandor, 2006

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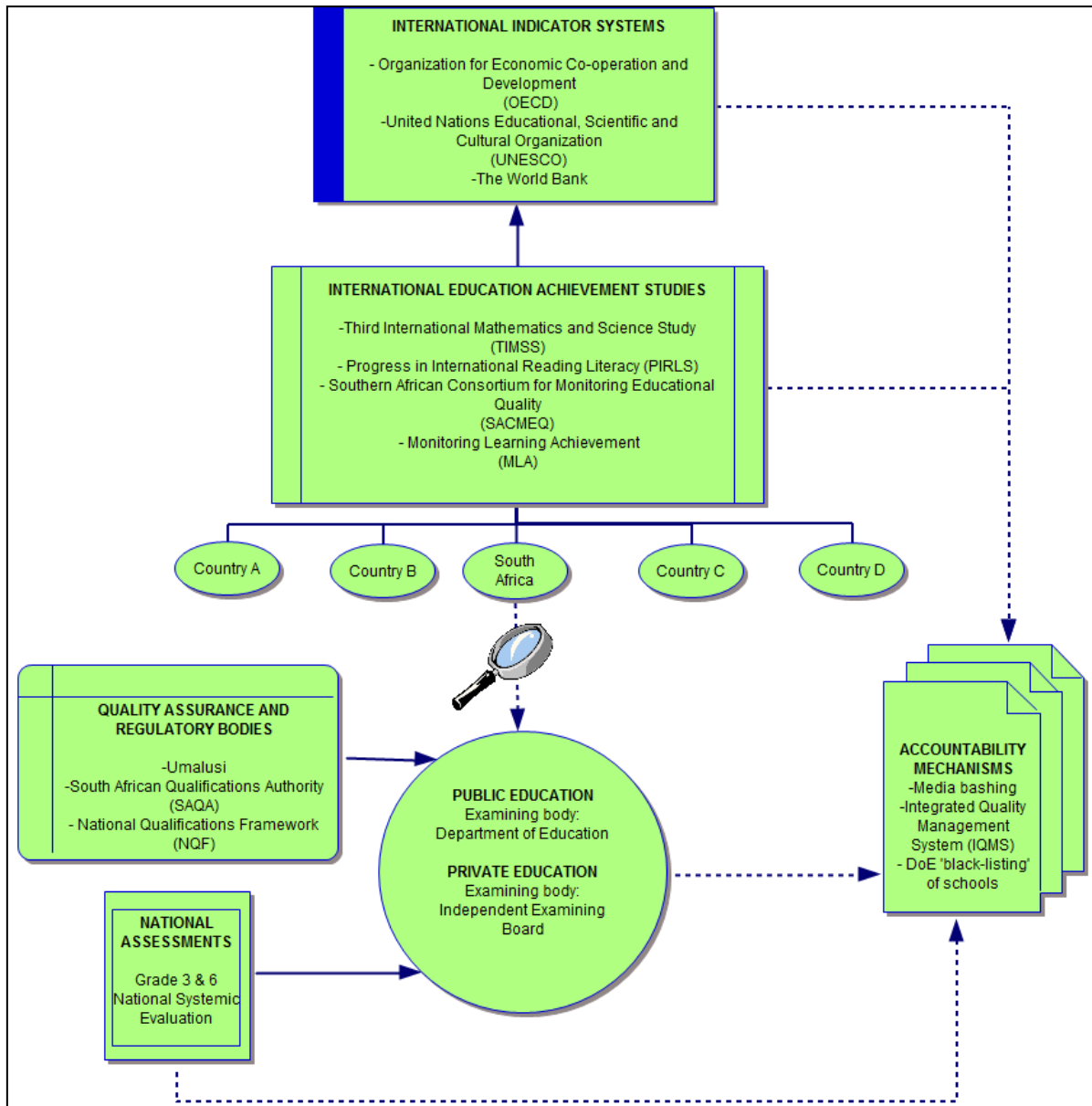
Figure 3.3: Some influences on the South African monitoring and feedback culture

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.
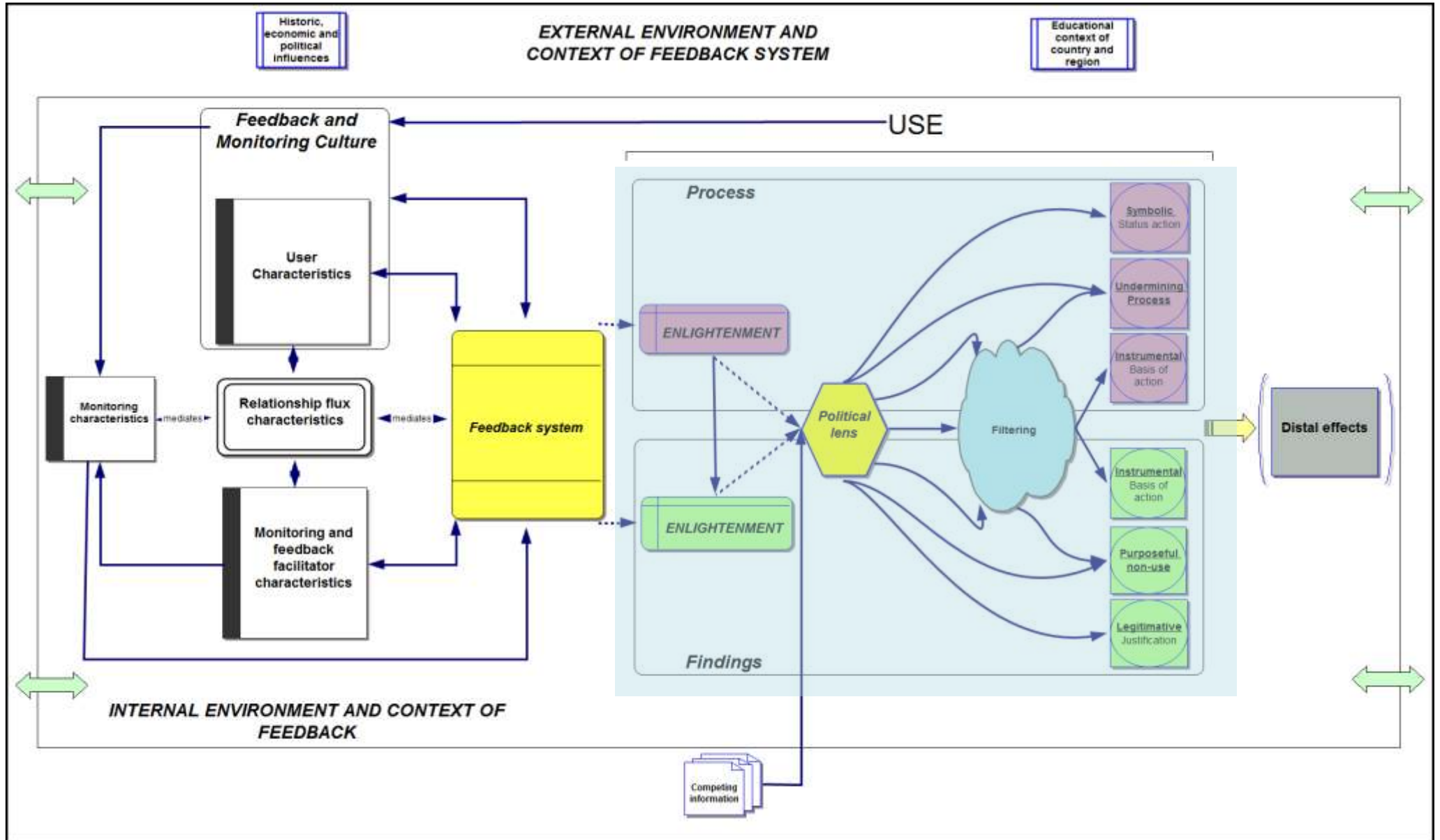
Figure 3.4: Preliminary conceptual framework for monitoring feedback use

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.
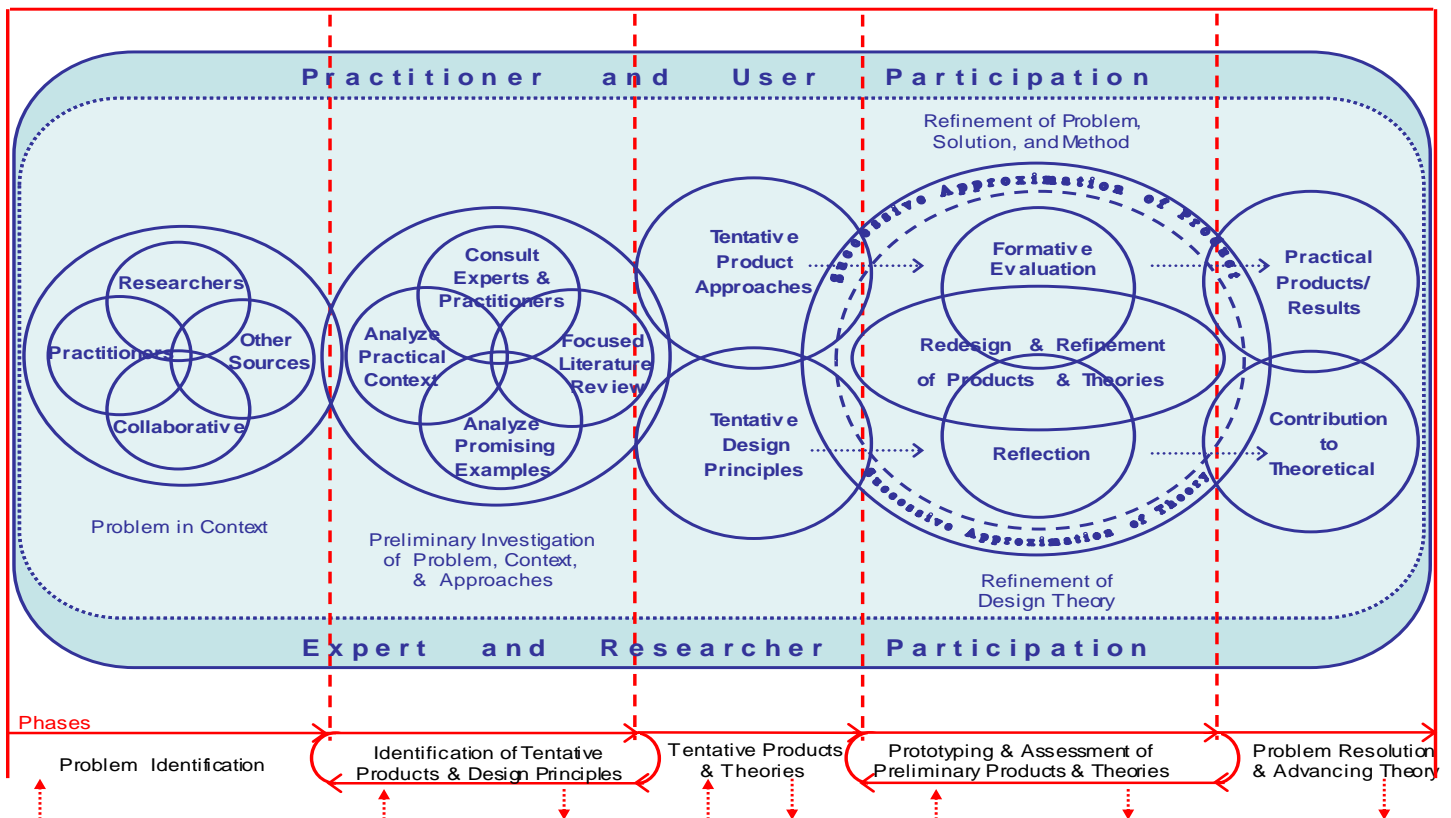

Figure 4.1: Wademan's Generic Design Research Model

---

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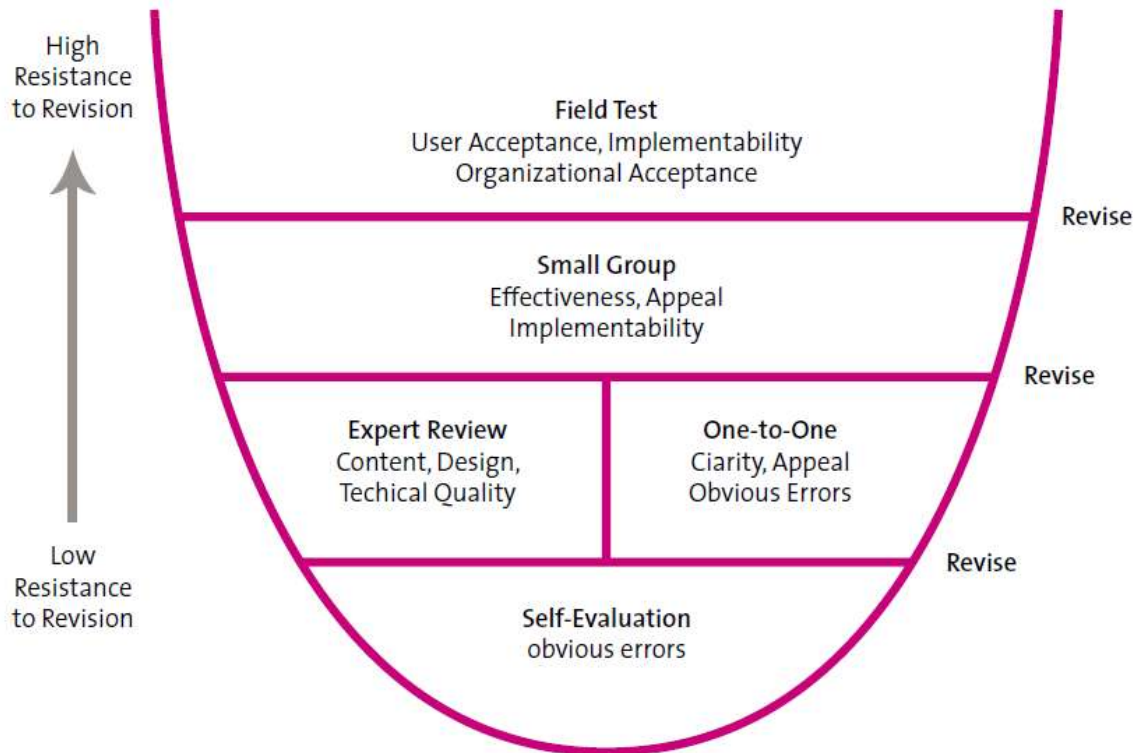
Figure 4.2: Layers of formative evaluation
(Tessmer in Plomp, 2009, p. 28)

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.
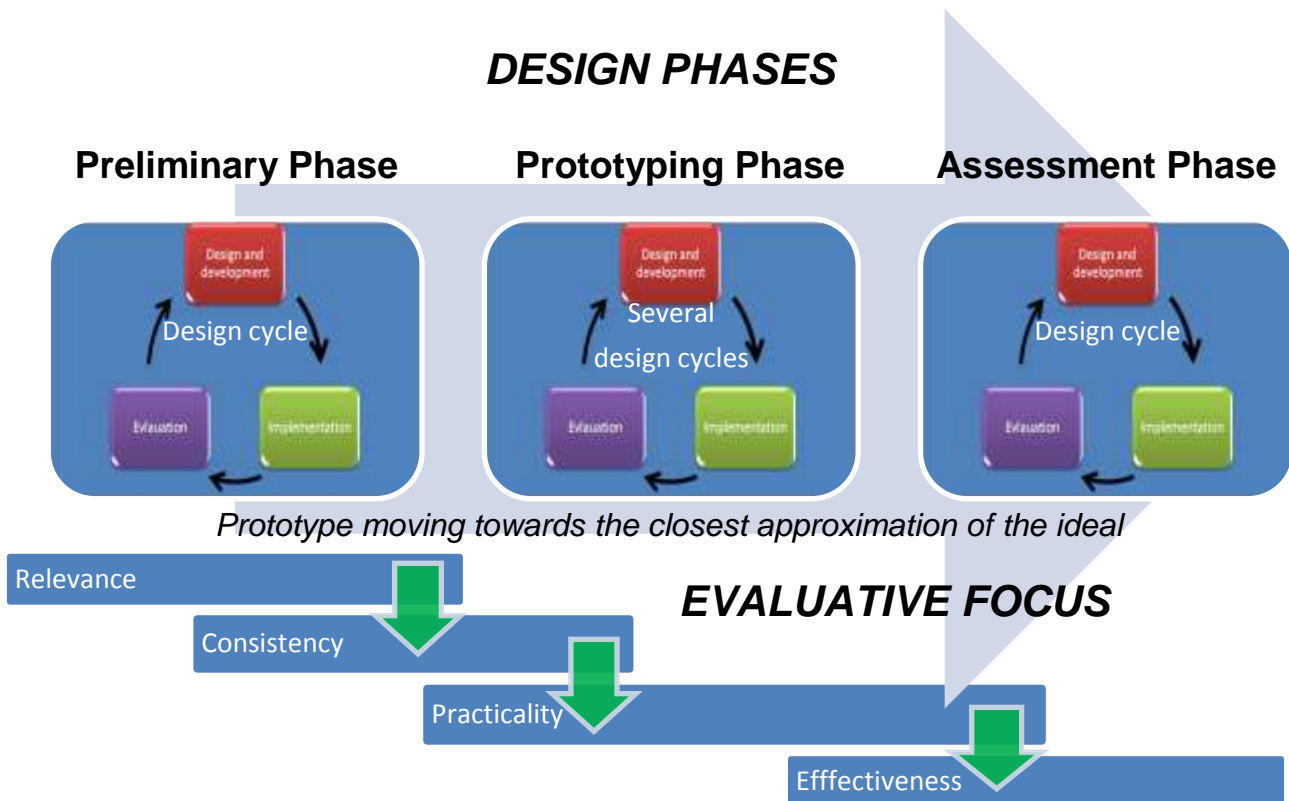
DESIGN PHASES

| Preliminary Phase | Prototyping Phase | Assessment Phase |

Prototype moving towards the closest approximation of the ideal

EVALUATIVE FOCUS

Relevance

Consistency

Practicality

Effectiveness

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.
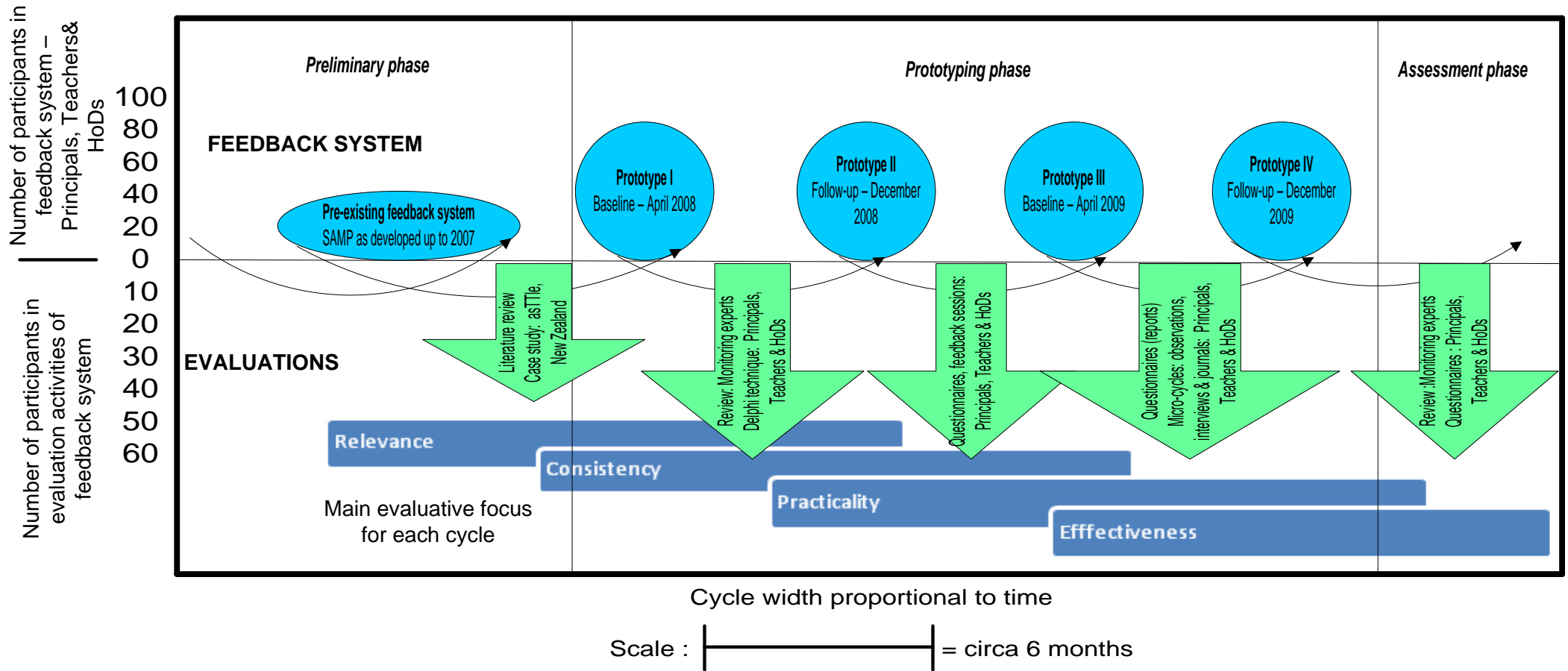
Figure 4.4: Design research cycles for this thesis

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.
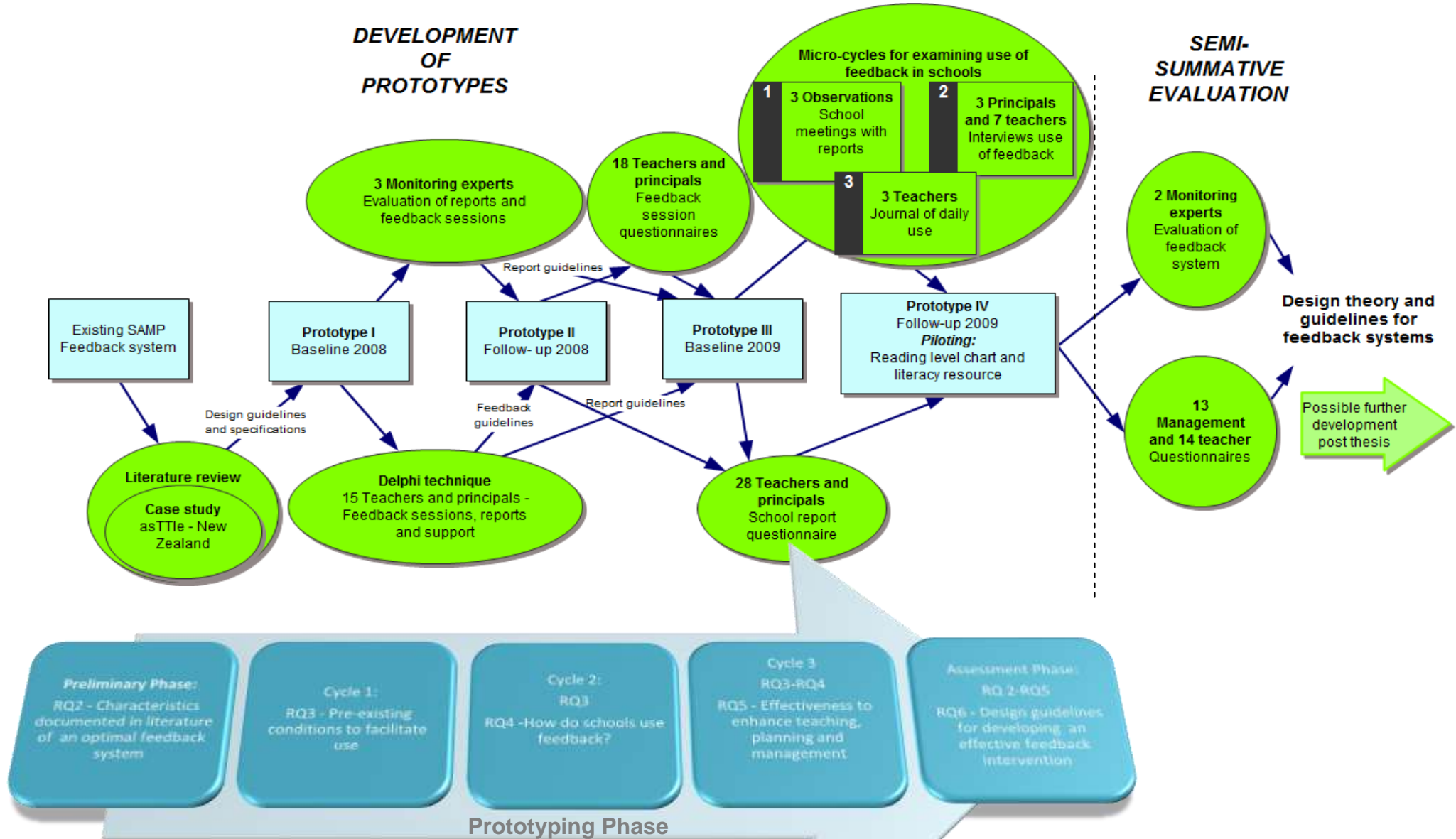
Figure 4.5: Design research process followed

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

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.