CHAPTER 8

FACTORS THAT INFLUENCE PERFORMANCE ON MIDYIS

The school system can be described as a nested system in which learners are situated within classes and classes within schools. Each level - learner, classroom, and school - interacts with each other in a way that results in a set of outcomes, in this case performance on the MidYIS assessment. This chapter provides an indication of the results of multilevel analyses undertaken. The aim is to explore which factors on a learner, classroom, and school-level influence the performance on the MidYIS assessment. Multilevel analysis was undertaken to shed light on which factors influences the performance of learners on the MidYIS assessment.

8.1 Introduction

> Even academic success which would appear to be clearly related to intelligence is affected by other factors: the skill of the teachers, the peer group of the children, the family circumstances and the health of the child (Kline, 2000, p. 33).

In Chapter 7 the relationship between MidYIS and academic achievement was explored. The results of the analyses show that a number of observations can be made about the variance explained across different schools. Although in some cases the percentage variance explained in MidYIS scale scores and the overall score is quite large, in the case of School 1, 66% of the variance in the total MidYIS score could be explained by achievement in English and 53% in mathematics (see Table 7.22 and Table 7.23). However, in other cases the percentage variance is small; for example in School 5 where 13% of the variance in the total MidYIS scores could be explained by achievement in English and 15% by achievement in mathematics. This leads one to the conclusion that aptitude or ability alone cannot account for the variance in academic achievement. According to Kline (2000), there is a correlation between ability, intelligence or aptitude and achievement due to common content of the

assessments and common skills they attempt to measure. However, this argument may be difficult to sustain in light of the nature of aptitude tests. Perhaps the answer lies in defining aptitude more generally as the ability to reason, as the ability to reason may be attributed to an inherited trait. Clearly other factors also play a part "…biometric research …demonstrate[s] that intelligence test scores are highly heritable… [but] not all variance is accounted for. What the environmental determinants are has yet to be determined empirically" (Kline, 2000, p. 82).

If the environmental determinants have not yet been empirically determined and if ability and skills can be taught, then the question is which school determinants may explain some of the variance unaccounted for? Learning and the development of skill take place within a school context and the context cannot be ignored (Luke, 2004). In education, learners are grouped together to form classes and classes collectively make up the learner body in a school. Thus the education system has a nested structure.

As was discussed in Chapter 2 and Chapter 3, school effectiveness models used to explore the effects of contextual factors at various levels of the system, are multilevel models. In a developing world context however, studies of school effectiveness seldom made use of the advanced statistical analyses such as multilevel analysis (Riddell, 1997). Not using this type of statistical analyses means that there is the risk of drawing inferences that are based on wrong assumptions, e.g. that effects at group level hold true for individuals (ecological fallacy) or that effects at the individual level hold true for the group level as well (atomistic fallacy) (Hox, 2002; Luke, 2004).

The aim of the present chapter is to explore factors on a learner, classroom, and school-level that have an effect on the overall performance on MidYIS, as was described in Chapter 5. The second main research question addressed by this exploration is ***which factors could have an effect on learner performance and therefore inform the design of the monitoring system?*** This broad research question comprises four specific research questions, as was discussed in Chapter 5:

2.1 What factors on a school-level affect the performance of learners on the assessment?
2.2 What factors on a classroom-level affect the performance of learners on the assessment?
2.3 What factors on a learner-level affect performance of learners on the assessment?
2.4 How can the identified factors be included in the design of the monitoring system?

Multilevel analysis is deemed appropriate to address these questions as it takes into account the nested structure of the education system. The variability in the upper levels of the nested system is also taken into consideration as the levels have an influence on each other. Additionally, the conceptual framework (see Chapter 3) underpinning this study consists of constructs operating at multiple levels within the school context (Luke, 2004).

Several issues pertaining to the multilevel analyses undertaken are addressed in this chapter. The data preparation, approach to model building, and the identification of possible factors to be included for analysis as was introduced in Chapter 5 are elaborated on in 8.2. This is followed by a discussion on the multilevel analysis undertaken (8.3). The null model is addressed first (8.3.1) followed by a discussion on the multilevel analysis using learner data, educator data and principal data (8.3.2). Concluding observations can be found in 8.4 of how the analyses address the four specific research questions in order to provide insights into the second main research question ***which factors could have an effect on learner performance and therefore inform the design of the monitoring system?***

8.2 Preparation for model building

Before applying multilevel analyses several steps must be taken, as was described in Chapter 5 (5.3.5.5). The data was first explored to ensure that assumptions underlying multilevel analysis was not violated, in particular multicollinearity. Multicollinearity exists when strong correlations (above 0.8) exist between two or more predictors in the model (Field, 2005). For the purposes of this exploration learners were linked to classes and classes to schools. In the section to follow the way in which variables were identified is elaborated on (8.2.1). This is followed by an overview of the approach to model building (8.2.2).

8.2.1 Identifying variables to be explored with multilevel analyses

Within the field of school effectiveness (see Chapter 2), factors associated with achievement are both broad and divergent, as the factors are operationalised differently across studies (Fertig, 2000). The contexts in which these studies took place also differ in terms of developing world contexts and developed world contexts. Studies from the developing world are characterised by large between school variation (Fertig, 2000) and challenges of studying classroom-level processes (Fertig, 2000; Scheerens, 2001a, 2001b), which is not necessarily the case with the developed world. In addition, research indicates stronger effects of material and human resource input factors in developing countries than in developed countries (Scheerens, 2001a).

In order to identify variables on the learner, classroom and school-level relationships were explored by means of correlation analyses between variables as taken from the questionnaires and the total score on MidYIS (see Table 8.1). It was found that the data was appropriate and that multicollinearity was not present (refer to Appendix L). In certain instances a variable was constructed out of a number of items - for example resources in the home corresponding with possible possessions in the home such as electricity, radio and television. However, in certain instances and based on literature, single items were used as a variable, such as level of education of mothers and fathers. In order to identify variables for further exploration all items, and where possible indicators based on a logical combination of items, were analysed. The criterion for inclusion for further analyses was based on the strength of the correlations (above 0.2) and their significance (0.99 confidence interval). The correlation analyses identified several moderate but significant relationships. In total six learner-level variables, six classroom-level variables and three school-level variables were identified. However, small sample sizes at the classroom and school-level was a concern and therefore only a limited number of variables could be included in the model, even though more variables were identified. A general rule of thumb of at least 10 observations per variable was used for analysis purposes (Field, 2005). The variables identified for inclusion were guided by prevalence in literature as well the strength and significance of correlations between the variables and the total score on the MidYIS assessment. Thus variables that were prevalent in literature and that had the strongest correlations were included.

Table 8.1 provides an overview of the factors at learner, classroom, and school-level included for further exploration.

Table 8.1 ***Description and significant correlations with MidYIS total of learner, classroom, and school-level factors***

Description of variables	Variable in MLwiN	Range	Correlation with the total MidYIS score
Resources in the home (composite variable $\alpha = 0.73$)	LeaResoHo	0-18	0.327**
With whom learners live	Lealive	1-4	-0.316**
Mother's education	Leamoted	1-5	0.382**
Father's education	Leafated	1-5	0.280**
Learners think it is important to do well in mathematics	Leamaimp	0-4	0.308**
Learners think it is important to do well in English	Leaengimp	0-4	0.363**
Lack of in-service training OBE is a challenge to assessment	Chalinservm (maths educator)	0-3	-0.360**
Resources available to educators (composite variable $\alpha = 0.95$)	Resoum (maths educator)	0-39	-0.367**
Mathematics teacher attitudes (composite variable $\alpha = 0.88$)	Teaattm (maths educator)	0-21	-0.463**
Lack of in-service training OBE is a challenge to assessment	Chaslinserve (language educator)	0-3	-0.316**
Resources available to educators (composite variable $\alpha = 0.93$)	Resoue (language educator)	0-39	-0.241**
Language teacher attitudes (composite variable $\alpha = 0.89$)	Teaatte (language educator)	0-21	-0.413**
Encouraging academic excellence	Prinencexc	1-4	-0.317**
Emphasis on achievement	Prinemach	1-5	-0.158**
Educators use monitoring systems in their classes	Prinedmon	1-5	0.301**

* α = *Cronbach Alpha*

** = *Statistically significant at 0.05 level*

Factors or variables having an impact on achievement are at the heart of the school effectiveness agenda. For the purposes of this exploration, achievement orientation and high expectations at a school-level have been included. These factors have a strong theory base (Bliss, 1991; Grey et al, 1999; Heck, 2000; Hill, 2001; Howie, 2002; Marsh, 1992; Newmann, 1991; Sammons, Thomas, Mortimore, Walker, Cairns & Bausor, 1998; Scheerens & Bosker,

1997; Scheerens, 1990, 1992, 2001a; Teddlie, 1994a, 1994c; Wills & Somers, 2001) and can be defined as having a clear focus on achievement and the mastering of subjects at a basic level in addition to encouraging high performance.

On a classroom-level, staff development or professional development has been identified as an important factor (Howie, 2002; Muijs, Harris & Chapman, 2004; Sammons, 1999; Teddlie, 1994b). The focus is on in-service training aimed at professional development in order to improve teaching practices (Halloway, 2003; Hirsh, 2005). Other factors included are resources available (Schereens, 2001a; 2001b) and educator attitudes.

Learner-level factors or variables include learner attitudes (Howie, 2002; Mortimore, 1998; Sammons, 1999), specifically towards English and mathematics. Depending on whether attitudes are positive or negative, behaviour may be promoted or inhibited in the classroom and at home (Anderson, 1994). Learner background characteristics are also included. The person(s) with whom the learner lives and the education of the mother provide some insight into the home environment of the learner and studies have linked these two factors to performance (Hortacsu, 1995; Milne & Plourde, 2006).

From the factors identified in Table 8.1 and the theoretical justification provided above, it is possible to construct a hypothetical model to be tested during the multilevel analyses. Figure 8.1 provides the proposed model. Three levels were identified. The school-level impacts on the classroom-level. The factors on the classroom-level as identified in Figure 8.1 impact the factors on the learner-level. However, it is also possible to assumed that perhaps the learner-level has a direct effect on the classroom-level and indirect effect on the school-level via the classroom-level while the classroom has a direct effect on the school-level.

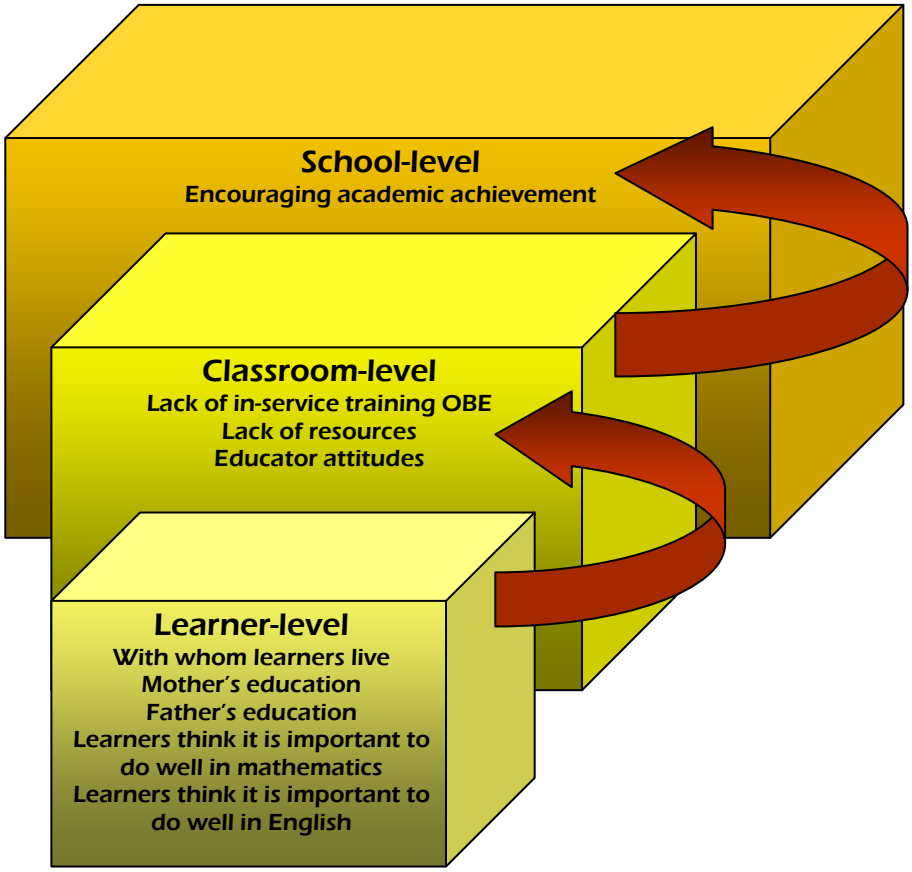

Figure 8.1 **Proposed model for the multilevel analyses**

8.2.2 Approach to model building

The three datasets, namely the learner, classroom educator and principal datasets, were merged into one dataset. Descriptive statistics were undertaken for identifying the mean, median and range in addition to identifying any missing values. The missing values were replaced with either the mean for the variable or the median as was discussed in Chapter 5. Ultimately 773 learners from 22 classes in 11 schools were included for analysis after replacements were made.

This research is exploratory in nature and the multilevel analyses progressed from the intercept-only or null model, to the final model (Luke, 2004). The model was built systematically by including variables on a one-by-one basis. Learner-level variables were added first, so that the contribution of each individual explanatory variable could be assessed (Hox, 2002). Each variable was added and analysed in order to ascertain whether the variable contributed to the model. This was done by identifying any change in the deviance;

here deviance refers to how well the model fits the data as was discussed in Chapter 5. Furthermore, whether the parameters were significant was also calculated by means of the Z-test, also known as the Wald test (Hox, 1995).

The intercept-only model gives an estimate for the intra-class correlation but also provides a measure of the degree of misfit in the model (Hox, 1995). After the intercept model had been examined and the intra-class correlations calculated, the first level or learner-level explanatory variables were added.

The parameters were fixed so that the contribution of each explanatory variable could be assessed. Here fixed means that the corresponding variance components of the slopes were set at zero (Hox, 1995). Full Maximum Likelihood (FML) estimation was used. This estimation method provides the opportunity to test the improvement of every consecutive model (Hox, 2002). This was done by means of computing the difference (chi-square variant) of the deviance from the model under investigation to the intercept-only model (Hox, 1995).

The second-level or classroom-level explanatory variables were then added and were evaluated. The classroom-level variables were explored in terms of random variance components. However, once random components had been introduced into the model non-convergence occurred. According to Hox (2002), the result of non-convergence is often an indication that random components can be omitted. Furthermore, sample size plays a large role as the sample size for this study was relatively small on the school and classroom-level (11 schools and 22 classes). The general rule of thumb is 30 groups and at least 30 individuals per group (Hox, 2002). Issues about sampling were elaborated on in Chapter 5. However, clearly due to the small sample sizes, methodological constraints were imposed, such as only including fixed parameters. Once significant variables were identified, chi-square analysis was undertaken in order to test whether the model in this step fits better than the previous model.

The third-level explanatory variables or school-level variables were added and the model examined to determine whether there were variables explaining between group variation (Hox, 2002; Luke, 2004). Once again only fixed parameters were included. Finally, cross-level interactions between explanatory group level variables and the individual level explanatory variables were explored. (Hox, 1995; Luke, 2004). However, for accurate and significant estimations, the number of groups should be larger than the number of individuals. Hox (2002) suggests that at least 50 groups with 20 individuals per group are needed (see

Chapter 5 for more details). Thus even though cross-level interactions were explored, the results were not significant.

8.3 The results of the multilevel analyses

Several models were run (see Appendix M for details). However for discussion purposes only the final models are presented in this section. The null model is discussed in 8.3.1. As was indicated in 8.2.2 this model contains no explanatory variables. Three models are discussed in 8.3.2. The models are the final models for the inclusion of explanatory variables at each level namely the learner-level, the learner and classroom-level and finally the learner, classroom and school-level.

8.3.1 The null model

The null model, or the intercept-only model, is the first step in building a multilevel model and does not contain any explanatory variables (Luke, 2004):

$$totalper_{ijk} = \beta_{0jk} + e_{ijk}$$

Where $\beta_{0jk} = \beta_0 + v_{0k} + u_{0jk}$

As was discussed in 8.2.3, the null model is an essential first step in the model building process as it provides a base from which consecutive models can be evaluated. The null model for this exploration (refer to Table 8.2) has an intercept of 47.995 (3.429) which is very similar to the overall mean for the sample (46.7%).

Table 8.2 ***The intercept-only model***

Effects	***Null Model***	
	Coefficient	**Standard error**
Fixed effects		
Intercept	*47.995*	*3.429*
Random effects		
σ_e^2	*129.120*	*6.664*
σ_{u0}^2	*11.997*	*6.752*
σ_{v0}^2	*121.412*	*55.146*
Deviance	**6013.45**	

The variance of the residual error for the learner-level is 129.120 (6.664), for the classroom-level 11.997 (6.752) and for the school-level 121.412 (55.146). The standard errors are all smaller than estimated parameters. The Wald test also referred to as the Z-test (Luke, 2004) was used as a significance test (Z=parameter/standard error of the parameter). This statistic is compared to a standard normal distribution. The aim is to test the null hypothesis that the parameter is zero (Hox, 2002). The result was statistically significant at $p<0.05$, indicating that effects do exist and that variables associated with the three levels should be included.

The intra-class correlations were calculated for both the classroom-level and school-level (see Table 8.3). The majority of the variance can be attributed to the learner-level which accounts for 49% of the total variance. Thus the remaining variance (51%) can therefore be attributed to the school and the classroom-level collectively. Of the 51%, 46% can be attributed to the school-level which is much higher than in developed countries (Luyten, personal communication, January, 2006). However, other studies undertaken in a developing world context confirms this result (Howie, 2002).

Table 8.3 ***Variance explained at the learner, classroom and school-level***

Level	Variance explained
Learner-level	49%
Classroom-level	5%
School-level	46%

8.3.2 The learner, classroom, school-level model

Several models were built in accordance with the procedure presented in 8.2.3 and discussed in detail in Chapter 5. Table 8.4 displays the results of three of the models in addition to the intercept-only model. The learner-level model was specified by the following equation:

$$totalper_{ijk} = \beta_{0jk} + \beta_1 lealive_{ijk} + \beta_2 leamoted_{ijk} + \beta_3 leamaimp_{ijk} + \beta_4 leaenimp_{ijk} + e_{ijk}$$

Where $\beta_{0jk} = \beta_0 + v_{0k} + u_{0jk}$

The result from the modeled equation (refer to Table 8.4) indicates that the four learner-level variables included in the model were all significant. Two learner-level variables namely resources in the home and father's education were included in previous models but were excluded from the final model as the variables were not significant nor did they substantially improve the fit of the model to the data.

The final model with the four variables did significantly differ from the null model indicating a good fit (difference in deviance 67.16). The model predicted that the score on MidYIS increases 1.175 percentage points when there are higher levels of education for mothers. Thus Learner A whose mother went to university could receive a total of 4.7 (1.175 x 4) percentage points more than Learner B whose mother has little or no formal education. Furthermore, it is predicted that learners who strongly agree that mathematics is important could score 5.98 (4 x 1.496) percentage points more than learners who do not agree that mathematics is important. A similar result emerges in terms of English. The model predicts that learners who strongly agreed with the statement that English is important receive 4.8 (4 X 1.189) percentage points more than learners who do not think that English is important. With whom the learner lives seems to negatively influence scores on the total MidYIS score. Thus it is predicted that learners who live with someone other than both of their parents or

guardians would score 4.11 (-1.371 x 3) less than learners who live with either their parents or guardians.

For the learner and classroom-level model, four learner-level explanatory variables and one classroom-level explanatory variable were included (refer to Table 8.4) in the final model for these two levels. In previously modeled equations resources and teacher attitudes for both mathematics educators and language educators in addition to challenge to assessment due to in-service training for language educators were included. However, these variables were not significant, had standard errors substantially larger than the estimated parameters and did not substantially improve the fit of the model (see Appendix M for examples). The final two-level model was specified by the following equation:

$$totalper_{ijk} = \beta_{0jk} + \beta_1 lealive_{ijk} + \beta_2 leamoted_{ijk} + \beta_3 leamaimp_{ijk} + \beta_4 leaenimp_{ijk} + \beta_5 chalinservm_{jk} + e_{ijk}$$

where $\beta_{0jk} = \beta_0 + v_{0k} + u_{0jk}$

The picture that emerged for the learner-level is very similar to what was described when the learner-level only model was discussed. Thus learners who live with either their parents or guardians scored 4 (1.349 x 3) more than learners who live with relative or some other arrangement. Similarly, learners whose mothers have higher levels of education scored 4.7 (1.186 x 4) percentage points more than learners whose mothers have lower-levels of education. In terms of the importance of mathematics and English, learners who agree that mathematics 5.97 (1.492 x 4) and 4.7 (1.176 x4) percentage points higher than learners who do not agree that mathematics, whilst for the importance of English the percentage points is 4.7 (1.176). Furthermore, learners who are taught by educators who feel that due to a lack of in-service training they are not able to use a variety of teaching and assessment methods as stipulated by OBE scored 8.23 (-2.745 x 3) percentage points less than learners who are taught by educators who feel that they are able to cope with OBE. However, with a standard error as large as the standard error for chalinservm (1.255) the result should be interpreted with caution (chalinservm - lack of in-service training OBE is a challenge to assessment for mathematics).

Table 8.4 ***Progression in model building***

Effects	Null model		Learner-level only		Learner and classroom-level		Final school classroom and learner-level	
	Coefficient	*Standard error*	*Coefficient*	*Standard error*	*Coefficient*	*Standard error*	*Coefficient*	*Standard error*
Fixed effects								
Intercept	***47.995***	***3.429***	***38.719***	***3.342***	***43.061***	***3.631***	***84.975***	***17.839***
Learner-level								
Lealive			-1.371**	0.380	-1.349*	0.380	-1.326**	0.379
Leamoted			1.175**	0.342	1.186*	0.342	1.189**	0.341
Leamaimp			1.496**	0.352	1.492**	0.352	1.467**	0.352
Leaengimp			1.189**	0.379	1.176*	0.379	1.140*	0.378
Classroom-level								
Chalinservm					-2.745	1.255	-2.677*	1.091
School-level								
Prinencexc							-18.991**	4.889
Prinedmon							8.878**	2.445
Random effects								
σ_e^2	129.120	6.664	119.207	6.153	119.226	6.154	119.156	6.150
σ_{u0}^2	11.997	6.752	8.819	5.254	8.023	4.931	8.246	5.018
σ_{v0}^2	121.412	55.146	80.356	36.957	57.041	26.780	15.158	9.286
Deviance	***6013.45***		***5946.29#***		***5942.147#***		***5929.887#***	

N=773 learners in 22 classes in 11 schools
** t-value > 2.58 a confidence interval of 99%
* t-value > 1.96 a confidence interval of 95%
# Deviance from null model to present model is significant at 0.01

In total, seven explanatory variables were included in the final model four learner-level variables, one classroom-level variable and two school-level variables. The final three-level model depicted in Table 8.4 is specified by the following equation:

$$totalper_{ijk} = \beta_{0jk} + \beta_1 lealive_{ijk} + \beta_2 leamoted_{ijk} + \beta_3 leamaimp_{ijk} + \beta_4 leaenimp_{ijk} + \beta_5 chalinservm_{jk} + \beta_6 prinencexc + \beta_7 prinedmon + e_{ijk}$$

Where $\beta_{0jk} = \beta_0 + v_{0k} + u_{0jk}$

The final model (refer to Table 8.4) is the best of the three models as it fits the data better, as indicated by the deviance, which is lowest of all the models. The model illustrates that if learners think that mathematics and English are important, live with either their parents or guardians and whose mothers have a higher level of education tend to score more percentage points. On the other hand, the result for learners who do not think mathematics and English are important, and/or do not live with either their parents or guardians and/or whose mother has little or no formal education is substantially lower. In terms of the second-level or classroom-level variable learners who are taught by educators who feel that their teaching practice is negatively affected by the lack of in-service training tend to score up to 8 (-2.677 x 3) percentage points less than learners who are taught by educators who do not share this view. Finally, it appears as if in schools where the principal does encourage academic excellence, learners tend to fare worse. It is possible that strategies and programmes are not put in place to add action to the vision of academic excellence. Alternatively it is possible that due to low morale among educators and learners academic excellence is not claimed as their own but rather externally enforced with little effect. In schools where principals indicated that educators do make use of monitoring systems, learners tended to fare better. Interestingly enough, when monitoring at the educator level is left out then emphasis on academic achievement is no longer significant. This indicates a relationship between these two variables.

Proportion of variance explained by consecutive models

> An important statistic…is the multiple correlation *R*, or the squared multiple correlation R^2 which is interpreted as the proportion of variance modeled by the explanatory variables (Hox, 2002, p. 63)

The proportion of variance modeled can be calculated by means of using the residual error variance, namely σ_e^2, σ_{u0}^2 and σ_{v0}^2, and the intercept-only model as a baseline (Hox, 2002).

In addition to calculating the proportion of variance explained (Table 8.5); the Akaike Information Criterion (AIC) was also calculated. The AIC is a fit statistic based on the deviance (Table 8.4) but also includes the number of parameters added (Luke, 2004). The

AIC was calculated by adding the deviance and twice the number of parameters. As with the deviance, the lower the AIC the better the model (Luke, 2004).

In the final model depicted in Table 8.5 the school-level variance is estimated at 87.5%, while on the classroom-level 31% is estimated with 7.7% estimated at the learner-level. Thus there is a higher proportion of variance explained *between* schools than *within* schools. When the learner-level model is considered 33.8% of the variance is explained between schools while only 7.7% can be attributed to the learner-level. As can be seen from the succession of each model the learner-level variance remains the same, which is to be expected. An interesting observation for the final model is that 87.5% of the variance is explained on the school-level. This is quite high, however the result may be explained by the fact that the schools were chosen according to maximum variation sampling (see Chapter 5) and there are only a small number of schools (eleven schools). However, clearly there are additional factors that would need to be explored at all levels to account for the unexplained variance.

Table 8.5 ***Proportion of variance explained by consecutive models for language***

	Model			
	Null	**Learner-level only**	**Learner and classroom-level**	**Final school classroom and learner-level**
School-level variance	0.46	0.338	0.53	0.875
	(46%)	(33.8%)	(53%)	(87.5)
Classroom-level variance	0.046	0.265	0.33	0.31
	(5%)	(26.5%)	(33%)	(31%)
Learner-level variance	0.49	0.077	0.077	0.077
	(49%)	(7.7%)	(7.7%)	(7.7%)
AIC	***6021.45***	***5962.29***	***5960.147***	***5951.887***

Even though the final model includes seven additional parameters when compared to the intercept-only model, this is still the best model. This model explains most of the variance and is the best model when the AIC statistic is considered, as the AIC is the smallest of the models explored.

Interaction effects

Three interaction effects were explored in this research, namely those between the school and classroom-level (principal encourages excellence and lack of in-service training) and between the school and learner-level (principal encourages excellence and mathematics is important, principal encourages excellence and English is important, challenge to assessment and educators make use of a monitoring system). However, no significant result was recorded within the framework of this study. This is perhaps not surprising as the sample sizes on the school and classroom-level are relatively small.

8.4 Conclusion

Learning can be influenced by a number of factors some of which are school and classroom related while others are not, for example, the environment in the home. However, if learning and achievement based on learning is to be understood, attempts should be made to explore the factors which impact on achievement. In this chapter, an attempt has been made to identify some of the factors which could have influenced the overall result on the MidYIS assessment. This influence, however, is not causal in nature but rather identifies tendencies. The exploration was guided by the main research question ***which factors could have an effect on learner performance and therefore inform the design of the monitoring system?*** This broad research question comprises four specific research questions. Each of the specific research questions are discussed separately in light of the findings presented in this chapter.

2.1 What factors on a school-level affect the performance of learners on the assessment?

Eighty-seven point five percent of the variance can be attributed to the school-level. This result is perhaps not surprising as other research from the developing world (Howie, 2002) has shown similar outcomes in terms of the large percentage of variance found at the school-level. It has to be kept in mind that the schools were selected by means of maximum variation sampling so schools which were vastly different were purposively selected and this could also account for the large percentage of variance. Two factors, of the three factors, on the school-level were included in this exploration, namely encouraging academic excellence and educators making use of monitoring systems. Academic expectations have to be translated into school policies and goals. Murphy (1988) reports that raising expectations and following this through with support programmes and staff development can increase the achievement of learners. Furthermore, in a study comparing high impact schools with average impact schools, it was found that high impact schools had a culture of high expectations. This culture of academic achievement was expressed in school policy documents and school practices focused on preparing learners for further education and the world of work (Perkins-Gough, 2006). Perhaps Murphy, Weil, Hallinger & Mitman, (1982, p. 24) said it the best "…schools that promote academic achievement have clearly defined goals based on academic matters."

Furthermore, monitoring of learner progress and indeed making use of monitoring systems has an affect on learner performance as substantiated in literature (Heck, 2000; Marsh, 1992;

Mortimore, 1998; Scheerens & Bosker, 1997; Scheerens & Creemers, 1999; Scheerens, 1992, 2001a; Teddlie, 1994a). A similar result was found in this study. This is elaborated on further in Chapter 9.

2.2 What factors on a classroom-level affect the performance of learners on the assessment?

Very often the aim of policy-makers and school management is to find the most effective remedy which will take the least amount of time to implement and which will be cost-effective. From a management perspective this makes sense but this does not make sense when whole generations of children are left behind because they cannot cope academically. Educators or teachers are essential if the success of learning and achievement as an outcome is to be ascertained, as stated by Bafumo (2005, p. 8) "… factors at the school, teacher and student level all impact on learning, but teachers are key to student achievement".

Six factors were identified to be included for exploration, namely resources, educator attitudes and challenge to assessment due to a lack of in-service training for both mathematics and language educators. Of the factors only one factor, namely challenge to assessment due to a lack of in-service training for mathematics educators, was included in the final model. Hirsh (2005, p. 38) concludes based on his research:

> …no single ingredient has greater impact on student achievement than the quality of the teacher in the classroom… not all teachers are adequately prepared to meet the diverse needs of today's students… Quality professional development employs these strategies, improves teaching, and closes achievement gaps.

Based on the quotation above, it may not be surprising that the issue of in-service training or rather the lack of in-service training is a prominent factor. If performance data is to be used by educators to focus on the specific needs of learners (Holloway, 2003), then educators need to know how to design effective assessments and use the information to guide their teaching practice. Furthermore, this factor alone accounted for most of the variance attributed to the classroom-level; this is discussed further in Chapter 9.

2.3 What factors on a learner-level affect performance of learners on the assessment?

Originally six factors were identified for exploration namely resources in the home, with whom the learners live, mother's education, father's education and the importance of mathematics and English. Of the six factors only four namely with whom the learners live, mother's education and the importance of mathematics and English, were significant. These four factors accounted for 7.7% of the variance. What seems to be clear is that the home environment of the learners has an effect on achievement. This result has been found elsewhere (such as the Coleman report released in 1966). For example in Nigeria specifically, the level of parental education, occupations of parents and size of family were correlated to achievement (Bolarin, 1992). More specifically perhaps, as in a study undertaken in Turkey, it was found that the mothers' level of education had a direct effect on learner performance (Hortacsu, 1995; Milne & Plourde, 2006). While the home environment seems to play an important role, learner attitudes and motivation seems to be important factors as well (Halawah, 2006; Howie, 2002).

2.4 How can the identified factors be included in the design of the monitoring system?

> Education is important. Educational policies need to be found that are effective and cost-effective. To achieve this demands that policies are based on sound evidence (Fitz-Gibbon, 2003, p. 313).

Clearly it is important to include the factors discussed in this exploration in a monitoring system using MidYIS. MidYIS was designed as a learner-level monitoring system. Thus it would seem plausible to include learner-level contextual factors. However, learning does not take place in a vacuum and as school effectiveness research has shown, factors on a classroom and school-level do have an effect on performance. For this reason, a monitoring system focusing on a single level has limitations. Perhaps a battery of instruments is required, in which instruments are associated with each level of the school system. Furthermore, in the context of South Africa and in light of the *Integrated Quality Management System (IQMS)* (as was discussed in Chapter 1) additional factors should be included so that schools can undertake self-evaluations. This is discussed in more detail in Chapter 9.

As has been illustrated in this chapter, factors at all three levels do account for the percentage of variance attributed to each level. However, there is still variance unaccounted

for in the models presented in this chapter, indicating that additional factors should be considered. Due to the practical constraint of sample size, it was not possible to do so. This does not mean that this is the end of the story. The exploration does lay the foundation for further analytical work to be undertaken. If the core factors based on sound empirical work can be identified then policy development and reform can take place. Furthermore, the value-added nature of the assessment when taken in conjunction with exit-level examinations provides additional information for use in self-evaluation exercises undertaken by schools.

CHAPTER 9

CONCLUSIONS AND RECOMMENDATIONS

As the final chapter of the dissertation, this chapter includes a summary and reflection on the findings. This is followed by reflections on the process, specifically in terms of the methodology used, situating this research within the field of school effectiveness and the contribution made to the broader body of knowledge. The reflections lead to a number of recommendations of what constitutes a suitable monitoring system for South Africa, how policy can be informed and what further research is needed. The chapter concludes with a discussion on the vital role feedback and intervention based on feedback plays in the utilisation of monitoring data.

9.1 Introduction

Learners may fail to reach their potential for a number of reasons that can be attributed to a range of social, school-based, and home-based factors. These may include large classes, inadequately trained educators, unsupportive educators and a school ethos based on academic competitiveness. For some learners, low levels of parental literacy may provide an additional barrier. In cases where parents are unable to give education-related help to their children at home, they may also lack confidence to approach the school should their children experience difficulties. Such parents are unlikely to provide a home environment to their children where literacy is valued (Hartley, 1990).

In order to be able to say anything about the performance of learners, the quality of the instrument used has to be considered (Luyten, Visscher & Witziers, 2005) and whether the instrument is fair to all learners has to be explored (Pelgrum, 1989). The aim of this research was to explore the possibility of using a monitoring system developed in the United Kingdom in the context of South Africa. What follows in this chapter is a reflection on the results of this research in an attempt to put forward recommendations on the use of monitoring systems for

practitioners and policy-makers alike. However, before presenting the recommendations, a summary of the results according to the research questions is given (9.2). This is followed by a discussion on and consideration of (9.3) the methodology used (9.3.1), reflections in light of school effectiveness research (9.3.2) and with how this research contributes to the body of knowledge in the domain of education (9.3.3). Recommendations are discussed in 9.4, specifically with regard to monitoring systems for South Africa (9.4.1), policy issues (9.4.2) and further research (9.4.3). The chapter is brought to a close by a discussion of the role of feedback and interventions in the utilisation of performance data received from monitoring systems (9.5).

9.2 Summary of the research

The issue of quality education is a topic of discussion, with South Africa facing the challenge of trying to implement policy on monitoring education. According to Pelgrum (1989), discussions about the quality of education occur in many societies with the aim to determine what learners learn when they are at school. Muller (2004, p. 221) states that assessment "is the most important system for signalling systemic efficiency and accountability". In South Africa, the use of assessment as an instrument to ascertain the efficiency of the education system began to enter into the "policy discourse" in the late 1990's (Muller, 2004, p. 224) and individual processes of evaluation were put in place (Muller, 2004). As was discussed in Chapter 1 the main policy foci were:

- Systemic Evaluation;
- Whole School Evaluation and more broadly perhaps;
- The Integrated Quality Management System.

What is clear is that there is a move to put policies in place to address issues of quality, equity, and redress. According to Muller (2004, p. 239)

> …we see a discernible move since 1994 away from an underdeveloped systemic policy (Grade 12 external assessment only) towards a marked progressive preference for formative, process, and integrative kinds of assessment with little real progress towards comprehensive systemic assessment.

Mechanisms for ascertaining the quality of education, in South Africa, are not functioning optimally in secondary schools, with the Department of Education mostly focusing its energies on the primary schools. What is clear is that without the necessary data provided by valid and reliable assessment instruments the "learning gaps in the system can't be known"

(Muller, 2004, p. 240). Without this information, informed decisions on interventions and indeed funding cannot be made (Muller, 2004).

The current research takes place against the backdrop of monitoring secondary education in order to ascertain the quality of teaching and learning. Monitoring, in this research, entails "not only the measurement of the output of a system, but also the evaluation of the measure" (Pelgrum, 1989, p. 8). This is by no means a small issue, as in the words of Sammons (2006, p. 2) "raising standards of achievement is seen as fundamental to economic performance and promotion of democratic engagement".

This research project is undertaken in collaboration with the Curriculum, Evaluation, and Management Centre (CEM) at Durham University in the United Kingdom and is funded by the South African National Research Foundation. The Middle Years Information System (MidYIS) project was originally developed by CEM with the aim of providing schools with information on how learners would perform at the end of two national examinations, namely Key Stage 3 and General Certificate in Secondary Education, in addition to providing value-added information. MidYIS makes use of a developed abilities assessment (see Chapter 4 for details). The assessment itself includes seven sub-tests that are combined to form four scales, namely:

1) Vocabulary scale
    - ❖ Vocabulary sub-test

2) Mathematics scale
    - ❖ Mathematics sub-test

3) Skills scale
    - ❖ Proof reading sub-test
    - ❖ Perceptual speed and accuracy sub-test

4) Non-verbal scale
    - ❖ Cross-sections sub-test
    - ❖ Block counting sub-test
    - ❖ Pictures sub-test

This research draws heavily on school effectiveness and school improvement literature as well as literature regarding the use of developed ability assessments (see Chapter 2). The central theme in school effectiveness research is the idea that schools do matter. The aim is

to disentangle the complex mix of learner characteristics and the educational experiences and to investigate how these interact to influence the development, progress and performance of learners (Sammons, 2006). The conceptual framework for this study drew on the work of Scheerens (1990). Scheerens (1990) developed an input-process-output model incorporating factors on the school and classroom-levels. This model was adapted and extended by means of including a learner-level and also by adding factors, which literature suggested as important for a developing world context.

Two main research questions were identified which can be divided into specific research questions and sub-research questions (see Chapter 3). These can be depicted graphically (refer to Figure 9.1).

The first research main research question guiding the study is ***how appropriate is the Middle Years Information System (MidYIS) as a monitoring system in the South African context.*** Here the word appropriate implies how suitable the MidYIS system would be for South Africa, looking specifically at issues related to how the MidYIS system compares with other monitoring systems, validity, reliability and what suggestions could be put forward so that MidYIS would be suitable for South Africa. Various facets of validity were investigated. In particular content-related validity (including curriculum validity), construct-related validity, and predictive validity were examined while inferences drawn with regard to reliability were done by means of internal consistency reliability. The first main research question has been operationalised by means of three specific research questions namely:

1.1. How does the Middle Years Information System (MidYIS) compare to other monitoring systems?

1.2. How valid and reliable are the data generated by the MidYIS monitoring system for South Africa?

1.3. What adaptations are needed to turn MidYIS into a monitoring system for the South African context

The second main research question extends the first research question. If MidYIS is valid, with South African adaptations, and reliable then what factors on a school, classroom, and learner-level could have an effect on learner performance. Thus the second main research question is ***which factors could have an effect on learner performance and therefore inform the design of the monitoring system.*** This research question has been operationalised by means of four specific research questions namely:

2.1. What factors on a school-level affect performance of learners on the assessment?

2.2. What factors on a classroom-level affect performance of learners on the assessment?

2.3. What factors on a learner-level affect performance of learners on the assessment?

2.4. How can the factors identified be included in the monitoring system?

A non-experimental pragmatic approach was adopted in this research (see Chapter 5). For pragmatism, both the meaning and the truth of any idea are functions of its practical outcome (Maxcy, 2003). It is the problem, which is of importance and not a preoccupation with methods (Creswell, 2003). Outcomes are what counts and not necessarily prior knowledge claims, laws or even what is true (Maxcy, 2003). Subjective and objective perspectives in addition to methods should be used in order to achieve the desired outcome. This integration of methods from the different paradigms is a powerful way of enhancing the credibility of findings (Petter & Gallivan, 2004). The view is held that there are similarities in the fundamental values between quantitative and qualitative approaches. These beliefs include the value-ladeness of inquiry, theory-ladeness of facts, that reality is multiple and constructed as well as that knowledge is fallible (Tashakkori & Teddlie, 1998).

Pragmatism lends itself to the use of mixed methods, which provides the researcher with the opportunity to answer the research questions adequately (Teddlie & Tashakkori, 2003). By using mixed methods, one may come to a more comprehensive understanding of the phenomena under investigation as this way one may develop a more complete portrayal of the social world as well as gain fresh perspectives and new ideas. The account of research using mixed methods is also more defensible as there is less bias as the one method compensates for the other method. Thus one is able to develop stronger knowledge claims (Greene, 2005). Mixed methods intentionally combine different tools and techniques to gather, structure, analyse and interpret quantitative and qualitative data (Williams, 1999). Various typologies can be identified under the banner mixed methods.

The typology used for this research is a concurrent nested strategy (refer to Chapter 5 for more detail). A concurrent nested strategy implies that there is a dominant method that guides the research. In the case of this research, a quantitative approach. The qualitative component was given lesser priority but was nested within the quantitative approach. The qualitative approach was embedded in the quantitative approach as the method addresses a different aspect of the question and seeks information from a different level. While the

quantitative approach in this study makes use of information at the school, classroom, and learner-level, the qualitative approach makes use of information at the provincial and national-levels.

Different data collection strategies were used in this research. In order to adequately address the research questions the following strategies were used (see Chapter 5):

- Curriculum document analysis (language and mathematics) was undertaken;
- Evaluation reports;
- Interview schedules were used;
- Questionnaires for the provincial officials, principals, educators and learners were utilised;
- A developed abilities assessment.

National Department of Education as well as Provincial Department of Education officials participated in this research. National officials in the field of assessment and curriculum were interviewed while the provincial officials in the fields of language and mathematics were asked to complete a questionnaire. One of the provincial officials was contacted telephonically and asked to elaborate on some of the answers provided in the questionnaire (refer to Chapter 5 and Chapter 6 for elaboration).

Apart from contacting Department of Education officials (both nationally and provincially), the assessment instrument was also sent for review. This review process fulfilled two purposes. Firstly, to ascertain what the overlap between the language and mathematics curriculum and skills assessed in the instrument would be. Secondly, to ascertain the correspondence of the sub-tests included in the MidYIS instrument and that of other developed abilities or aptitude tests. The review was undertaken by language and mathematics specialists as well as educational and research psychologists. By undertaking a thorough analysis of the language and mathematics curriculum documents, depth was added to the evaluation process (refer to Chapter 5 and Chapter 6 for further elaboration).

In addition to the Department of Education officials and specialists in the field of psychology and education, eleven secondary schools in the Pretoria area also participated in this research. The Department of Education officials were purposefully selected. The eleven schools were sampled by means of maximum variation sampling so that schools selected would be representative of the different types of schools across South Africa. Two Grade 8 classes were randomly selected from each of the schools. The principal of each of the

schools as well as the language and mathematics educators of the two classes selected were asked to completed questionnaires (refer to Chapters 5, 7 and 8).

Figure 9.1 provides a diagrammatic view of the research questions used to guide this research.

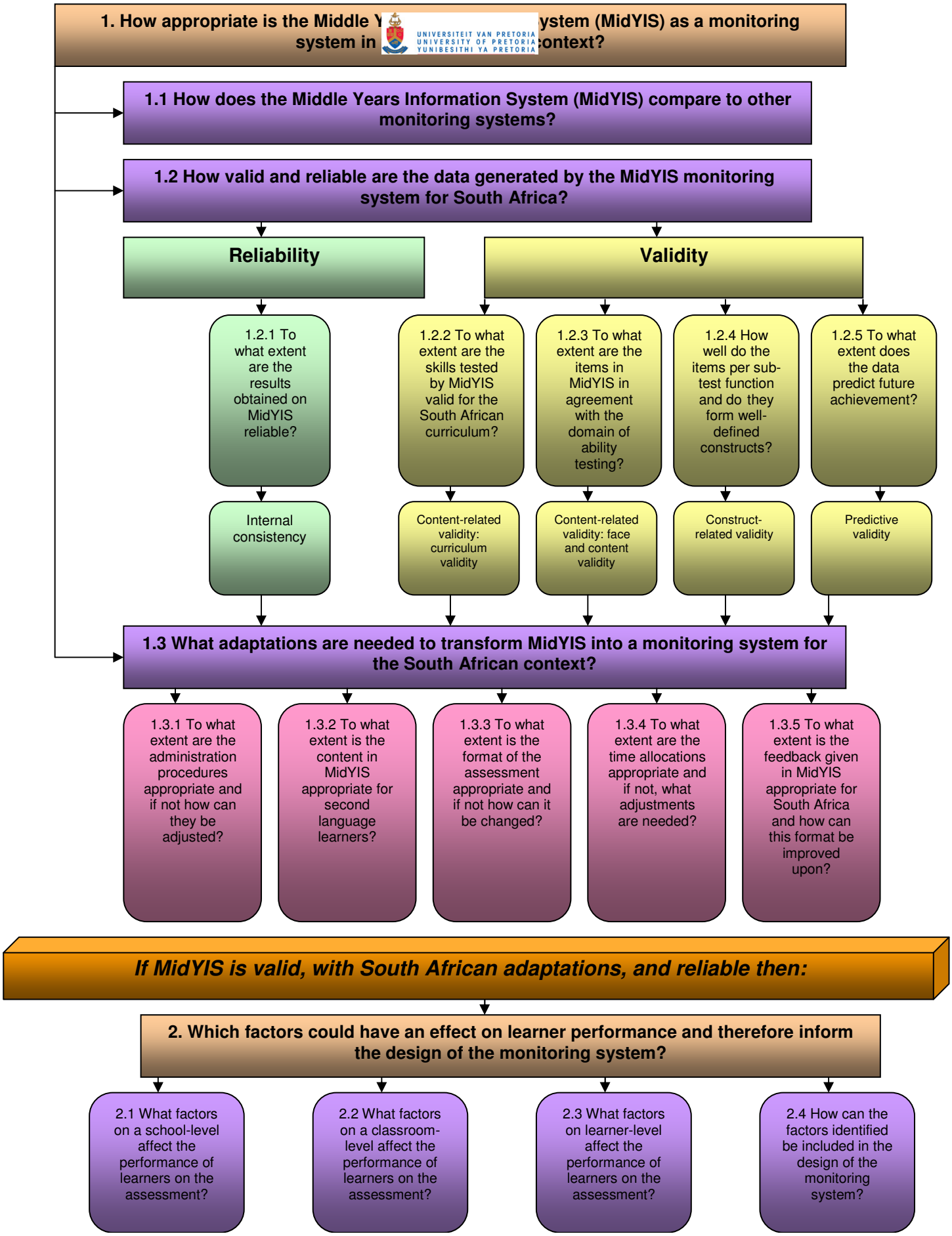

Figure 9.1 **Overview of the main research, specific research and sub-research questions**

The first specific question identified for the first main research question is ***how does the Middle Years Information System (MidYIS) compare to other monitoring systems?*** This question was addressed in Chapter 2 and Chapter 4 and is a reflection on insights drawn from the literature review. Table 9.1 provides an overview of three monitoring systems that were reviewed in conjunction with MidYIS, namely the ZEBO-project, the VCE data project, and the ABC+ model. What is striking from this comparison is that MidYIS essentially only includes the learner-level while the other systems include at least two levels (learner and classroom or learner and school). Furthermore, while MidYIS does perform a monitoring function, its main aim is to provide schools with value-added information on performance and the schools decide how to use the information. CEM processes the information and distributes the data in user-friendly form for the school managers and educators to analyse further. Although summaries are provided by CEM, the schools are responsible for interpreting the data and undertaking additional analysis. This seems to be a key point of all the systems included in Table 9.1 in addition to the idea that the systems should not be too intrusive on school time. A divergent point, however, is the inclusion and use of behavioural components such as learner attitudes. In MidYIS this is available but schools decide whether they want this additional information. While it would appear that the other systems include behavioural information as an integral part of the monitoring system, and not an additional component as with MidYIS.

The MidYIS system as discussed in Chapter 4 makes use of an abilities assessment and not a curriculum-based assessment. The ZEBO-project in the Netherlands comprises a curriculum assessment, abilities assessment and background questionnaires. This is similar to the VCE data project where both curriculum-based and abilities assessments are used. For South Africa it may be beneficial to include under the banner of South African Secondary School Information System (SASSIS) a suite of instruments. On a learner-level a curriculum-based assessment and an abilities assessment should be included in addition to a questionnaire which would provide background and attitudinal information. This would be in line with monitoring systems in the developed world. Furthermore, a national examination should also be used in order to provide additional value-added information and to explore predictive validity. In this regard, exit-level examinations at Grade 9 and Grade 12 would be appropriate. A questionnaire, classroom observations using defined protocols and perhaps follow-up interviews would be appropriate for the classroom-level while on the school-level a questionnaire and follow-up interview could be included.

What the monitoring projects indicated in Table 9.1 do not include are levels other than those directly related to the school. For South Africa it would be beneficial if the district and

provincial levels were included in a newly developed monitoring system. On the one hand additional information relating to the support given to schools can be ascertained. On the other hand the data collected on the school-level can be processed in a manner which would facilitate the development and implementation of additional intervention programmes at the school, district and/or provincial-level if needed. The support from the district-level given to schools in addition to intervention programmes developed by the educators within the schools could potentially be the difference of success or failure of the intervention programme.

Table 9.1 ***A comparison of the ZEBO-project, VCE data project, ABC+ model and MidYIS***

System Characteristics	The ZEBO-project (The Netherlands)	The VCE data project (Australia)	The ABC+ model (The United States of America)	Middle Years Information System (The United Kingdom)
Unit of analysis	School, classroom, and learner-level.	School, classroom, and learner-level.	School, classroom, learner, and parent level.	Learner-level.
Rationale underpinning the project	Developing sound self-evaluation tools based on research and theory.	Assist schools to monitor the effectiveness of their teaching and learning.	To provide process information which schools can use for improvement plans.	To provide schools with value-added information.
Stakeholder input	Schools evaluate themselves. Component evaluated to ascertain, efficiency, effectiveness and use of information.	Schools interpret the data based on training received. School management teams primarily responsible. However, the process is participative and the stakeholders work together.	Stakeholders decide which elements should be monitored and who will collect the data Participative in nature.	Schools interpret the data based on training received.
Effect on behavioural aspects	Information used by schools to draw up self-improvement plans in line with legislation.	Information used by schools to develop strategies for improvement including personnel management strategies.	Information used to develop school improvement strategies and plans.	Schools decide whether they want information on behavioural aspects.
Implementation of the project	School-based minimum interference with school activities.	Minimum interference with school activities as this forms part of the VCE assessment programme.	The model is time-consuming and labour intensive. However, data collected is not collected by outcomes-driven indicator systems.	School-based minimum interference with school activities.

The second specific question for the first main research question was ***how valid and reliable are the data generated by the MidYIS monitoring system for South Africa?***

This specific research question comprised several sub-questions:

1.2.1. To what extent are the results obtained on MidYIS reliable?

1.2.2. To what extent are the skills tested by MidYIS valid for the South African curriculum?

1.2.3. To what extent are the items in MidYIS in agreement with the domain of ability testing?

1.2.4. How well do the items per sub-test function and do they form well-defined constructs?

1.2.5. To what extent does the data predict future achievement?

Different strategies for making inferences related to validity were presented in Chapter 5 ranging from conceptual considerations as is the case with content-related validity, (presented in Chapter 6), to empirical considerations as is the case on construct-related validity and predictive validity (presented Chapter 7).

The sub-question 1.2.1 is related to the reliability of the MidYIS results namely ***to what extent are the results obtained on MidYIS reliable?*** The reliability of the assessment instrument is addressed in Chapter 7. The analysis was undertaken with the whole sample and with learners from different population groups. Although initial results indicate internal consistency and that items do measure the same construct, larger samples than those included in this study for sub-population groups would be required if inferences per population group were to be made with more confidence.

Chapter 6 of the dissertation focused on issues associated with the content-related validity of the MidYIS assessment, namely sub-questions 1.2.2 and 1.2.3. As deduced from the two sub-research questions, the content-related validity of MidYIS can be evaluated from two perspectives, namely a curriculum perspective and a psychometric perspective. Although these two perspectives are addressed separately, there is an apparent link between them. From a psychometric perspective, MidYIS is a developed abilities assessment. Ability is a competence, a skill or an aptitude and the curriculum can have its roots in competency-based education, as is the case in South Africa. Due to this interrelatedness of MidYIS as a developed abilities assessment and the South African curriculum with its roots in competency-based education, both aspects had to be explored.

The sub-question 1.2.2 or ***to what extent are the skills tested by MidYIS valid for the South African curriculum*** was explored by means of curriculum document analysis and specialist evaluations, while background information was provided by the National and Provincial Department of Education. The clear message from the National and Provincial Departments of Education was that any assessment used in a school setting must be aligned to the curriculum. In order to explore the alignment of the MidYIS assessment with the South African curriculum, document analysis was undertaken and specialists consulted. Two learning areas were selected, namely language and mathematics, as the fundamental skills assessed in MidYIS corresponded with these two learning areas (refer to Chapter 5).

Three of the six outcomes in the language learning area were represented in the MidYIS assessment indicating a moderate alignment between MidYIS and the South African curriculum (refer to Chapter 6). For the language learning area three of the six outcomes are not represented. However, the skills assessed in the MidYIS assessment which can be found in the curriculum refer to the basic skills needed, for example skimming, scanning, punctuation and vocabulary. It is clear that even though the MidYIS assessment does not directly include three of the six learning outcomes, what it does include is the basic skill that is needed to succeed in the other learning outcomes included in the language learning area. However, it is possible to construct additional scales that directly relate to the other learning outcomes, such as reading a passage and answering questions related to the passage. By means of including an additional section, learner reading skills and comprehension can be directly assessed.

Inferences, in terms of curriculum validity for the mathematics learning area, are substantially stronger; as four of the five learning outcomes are represented in MidYIS (refer to Chapter 6). It would appear from the document analysis and specialist evaluation that MidYIS has a high degree of curriculum validity, especially for mathematics. However, additional items pertaining to the outcome currently not represented, namely data handling, may make inferences stronger.

The sub-question 1.2.3 focuses on content-related validity was ***to what extent are the items in MidYIS in agreement with the domain of ability testing and applicable for South Africa***. This question was addressed via expert appraisal. The experts were selected from fields of educational and research psychology. The evaluations from the psychologists indicate that the items in the MidYIS are in agreement with the ability domain. Furthermore, MidYIS is comparable to other ability assessments currently used in South Africa such as the

Differential Aptitude Test (DAT) and is not biased in terms of gender or race (refer to Chapter 7).

The sub-question (1.2.4) ***how well do the items per sub-test function and do they form well-defined constructs*** was addressed by means of item and scale analyses (as described in Chapter 7), specifically Rasch analysis for item level analysis. What emerges from the Rasch analysis is that there are core items associated with sub-tests and that the sub-tests can be integrated into scales, as was originally designed by CEM. However, there are items which seem to be measuring constructs other than those they were intended to measure (see Chapter 7 for details). Thus the items which were identified as misfitting should be revised or rewritten based on an assessment framework for the assessment as a whole. The assessment framework should be developed from both a curriculum and psychometric perspective. An explanatory note of the fitting or rather misfitting of items or persons is needed. In Rasch analysis, fit is not interpreted in the same way as in the world of measurement where one would state that the model fits the data. Rather, fit statistics are used to detect discrepancies between the Rasch model prescriptions and the data (Bond & Fox, 2001). Misfitting persons, in Rasch analysis, represents the degree to which the response pattern of the individual is more haphazard than the Rasch model would have expected. The unexpected response pattern could indicate more or less variation than expected.

The sub-question 1.2.5 is related to the predictive validity of the assessment, namely ***to what extent does the data predict future achievement?*** The analyses were undertaken per school and not across schools as standardized national examination or other assessment results were not available and therefore school-based results were used. The results indicated that the scales as constructed using the Rasch analysis do correlate with the results obtained from schools, with most of the correlations above the 0.3 criterion stipulated by Kline (1993). So MidYIS could possibly be used for prediction purposes in the context of South Africa. However, further analytic work is needed before definite inferences can be drawn. It would be appropriate to increase the sample to included schools from different provinces (including rural schools) as well as to use a standardised national school-based assessment in this regard. What seems to emerge is that MidYIS on its own can only account for a certain amount of variance and there are other factors on the learner, classroom and school-level that have to taken into account (see Chapter 7 for details).
It was clear that adaptations had to be made to MidYIS to make it relevant for South Africa (see Chapter 6). Some of the adaptations were easier to effect than others. The adaptations needed to range from allocating more time per sub-test to possibly including new items to

existing sub-tests or adding additional sub-tests to the assessment. The specific research question of ***what adaptations are needed to transform MidYIS into a monitoring system for the South African context*** (1.3) was addressed based on the reports of specialists in the fields of language and mathematics. The specialists in the fields of language and mathematics suggested that the administration procedures be reviewed, the appropriateness for second language learners be established, the format be reviewed, the time limits be evaluated, and the way in which feedback is given through to schools be assessed.

Sub-research question 1.3.1 ***to what extent are the administration procedures appropriate and if not, how can they be adjusted*** was explored by means of expert appraisal. The expert evaluation reports indicated that the instructions were ambiguous and could be difficult to follow. Thus the instructions were revised, based on the suggestions provided by the specialists, so that learners would understand what was expected of them but so that the revised version would still be comparable to the original (refer to Chapter 6 for details).

In answer to sub-research question 1.3.2 ***to what extent is the content in MidYIS appropriate for second language learners***, the experts indicated that a number of items would not be accessible for second language learners. The specialists identified these items and also provided feasible alternatives. The changes suggested by the specialists were effected (see Chapter 6).

Sub-question 1.3.3 ***to what extent is the format of the assessment appropriate and if not, how can it be changed*** was also explored by expert appraisal. Overall the format of MidYIS was acceptable. However, the specialists indicated that should a learner be unsure of what to do they would have to page to the beginning of the sub-test in order to reread the instructions. This wastes time. Therefore, the instructions were included at the top of the page throughout MidYIS, as suggested by the specialists, so that learners if uncertain could reread the instructions without wasting time (refer to Chapter 6).

To what extent are the time allocations appropriate and if not, what adjustments are needed is sub-research question 1.3.4. The experts were not happy with the time limits allocated for various sections of MidYIS. Therefore, the time allocated for each sub-test was increased based on the recommendations of the specialists so that the majority of the learners would be able to complete or almost complete the sub-test. This is also in accordance with the type of assessment, as MidYIS is a combination of a speed and power test as was discussed in Chapter 5 (see Chapter 6 for an elaboration).

The final sub-question 1.3.5 is ***to what extent is the feedback given in MidYIS appropriate for South Africa and how can this format be improved upon?*** Although this question was not directly addressed in this research, recommendations can be made on the basis of literature. The ultimate use of assessment information is that it is elicited with the goal of improving teaching and learning. According to Van Petegem, Vanhoof, Daems and Mahieu (2005) there are four reasons to gather performance data, namely for information needs, for accountability purposes, creating marketing mechanisms or to stimulate discussions on quality in education. An essential component in all of these reasons is the way in which the performance information is provided (Vanhoof & Van Petegem, 2005, p. 206):

> 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.

Feedback, according to Black and Wiliam (1998), should be about particular qualities of learners and learners' work and how the learner can improve. If monitoring systems are to provide the information needed to assist schools then the research agenda has to be guided by the following questions (Luyten, et al., 2005):

- How can the feedback be made accessible? Currently, the MidYIS system provides the minimum information to schools and the information is illustrated in the form of tables and graphs. School management teams are provided with training that equips them to undertake further analysis of the data. In South Africa the approach used by CEM, although cost efficient, would not work, as there are not many schools in a position to pay for the services of the CEA. Educators and schools need to have the information presented to them in a way that is easy to understand and recommendations given should be based on the results. Not only should the feedback be made easily accessible, it should be followed up by a support component.
- What information is deemed credible by schools? Schools should provide an indication of what type of information is needed. For example, it is plausible that schools may be more interested in academic achievement than the learners' perception of school climate. A collaborative partnership between the schools and CEA should be developed in order to ascertain what information is needed.
- What type of feedback is most accessible and easy to understand? Do educators and school managers prefer graphical representations, narrative descriptions or tables? School managers and educators are the experts in their fields and should not be

patronised. Therefore, the form of the feedback should be formulated by the stakeholders. This means the schools would be more likely to use the information if it is presented in a manner recommended by them.

- How can feedback systems be used to detect problems and find solutions? The type of feedback given should provide an overview in addition to potential diagnostic information. The aim of any monitoring system is to identify problem areas and to develop an intervention to address the problems. This should be done in collaboration with the stakeholders to ensure that ownership of the process is taken.
- What strategies for change are most effective? This is an important component and addresses the question of what worked and what did not.

Learning can be influenced by a number of factors, some of which are school related and others are not. However, if learning and achievement based on learning is to be understood, attempts should be made to explore the factors, which impact, on achievement. In Chapter 8, an attempt was made to identify some of the factors, which could have influenced the overall result on the MidYIS assessment. The exploration was guided by the second main research question ***which factors could have an effect on learner performance and therefore inform the design of the monitoring system?*** This broad research question comprised four specific research questions. Each of the specific research questions is discussed separately in light of the findings presented in Chapter 8. This is an extension of the first main research question (***how appropriate is the Middle Years Information System (MidYIS) as a monitoring system in the South African context***) and makes use of a multilevel model in order to provide some insights into the specific research questions (see Chapter 8). The data were explored to ensure that the assumptions of the statistical analysis were not violated. Specifically, mulitcollinearity was investigated. It was found that no assumptions were violated.

The first specific research question (2.1) is ***what factors on a school-level affect the performance of learners on the assessment?*** The multilevel analysis showed that 85.7 percent of the variance could be attributed to the school-level. This result is perhaps not surprising as other research from the developing world (Howie, 2002) has shown similar outcomes in terms of the large percentage of variance found at the school-level. Two factors (of the three factors) on the school-level were significant, namely encouraging academic excellence (negative effect) and educators make use of monitoring systems (positive effect). Academic expectations have to be translated into policies and goals. Perhaps Murphy et al. (1982, p. 24) capture this idea the best when they say "…schools that promote academic achievement have clearly defined goals based on academic matters". The negative effect of

academic performance is a surprising result. It is possible that although the principals indicated that they do encourage academic achievement something else is happening in the school or home environment that is not translated into the results that the school would want.

The second specific research question (2.2) is ***what factors on a classroom-level affect the performance of learners on the assessment?*** Six factors were identified for exploration (three for the language educator and three for the mathematics educator), namely resources, educator attitudes and challenge to assessment due to a lack of in-service training. Of the factors only one factor namely, challenges to assessment due to a lack of in-service training (negative effect) for the mathematics educator was included in the final model. This factor alone accounted for 31% of the variance for the classroom-level. This could be a consequence of small sample sizes but the result is significant nonetheless. This is perhaps not surprising because if performance data is to be used by educators to focus on the specific needs of learners (Holloway, 2003) then educators need to know how to design effective assessments and use the information to guide their teaching practice.

The third specific research question (2.3) ***what factors on a learner-level affect performance of learners on the assessment?*** Originally six factors were identified for exploration namely resources in the home, with whom the learners live, mother's education, father's education and the importance of mathematics and English. Only four of the six factors were significant, namely with whom the learners live (negative effect), mother's education (positive effect) and the importance of mathematics and English (positive effect). These four factors accounted for 7.7% of the variance. Although these four factors combined accounted for only a small percentage of variance, their inclusion dramatically improved the fit of the model to the data.

The final specific research question (2.4) is ***how can the factors identified be included in the design of the monitoring system?*** Clearly the factors, in this exploration, are important to include as part of a monitoring system using MidYIS. MidYIS was designed as a learner-level monitoring system. Thus it would seem plausible to include learner-level contextual factors. However, learning does not take place in a vacuum and as school effectiveness research has shown a number of factors on a classroom and school-level do have an effect on performance (Sammons, 2006). Monitoring systems generally do include at least two levels as was discussed in Chapter 2. A monitoring system focusing on a single level may have some drawbacks such as not targeting higher-level variables to monitor change, although the system itself is less complex. It has been suggested that a battery of instruments should be called for. This implies that instruments are associated with each level of the school system. This is discussed further in the recommendation section to follow.

However, the choice of which level to target or which levels to include will be determined by educational policy makers and practitioners. It may be that policy-makers or practitioners may be interested in only one level. So it may be that only learner performance and limited number background indicators are of importance. At the same time, it may be of interest to include indicators on the classroom or school-level in order to obtain a complete picture that could be targeted by intervention programmes.

9.3 Discussion and reflection

There are several areas of reflection which are worth discussing. These reflections are elaborated on in the section to follow and centre on methodological reflections (9.3.1), the field of school effectiveness research (9.3.2), and finally how this research contributes to the body of knowledge, practical and scientific (9.3.3).

9.3.1 Methodological reflections

In this research, mixed methods were applied. A continuing issue in the mixed methods discourse, however, is the manner in which paradigms are used in the development of the mixed methods as a field (Teddlie & Tashakkori, 2003). Based on the pragmatic paradigm the use of mixed methods is appropriate in the case of this research and the use of the methods is seen as being complementary (Morse, 2003), as one method is only a partial snapshot of the phenomenon and the use of both methods provides a more complete picture (Greene & Caracelli, 2003). Much work needs to be done in the area of mixed methods research regarding its philosophical underpinnings, designs and data analysis, validity strategies as well as rationale for mixing and integrating procedures (Johnson & Onwuegbuzie, 2004).

This research was primarily exploratory in nature as is reflected in the research design, such as the sample of respondents, selection of data collection methods and analysis techniques. The first main research question ***how appropriate is the Middle Years Information System (MidYIS) as a monitoring system in the South African context*** does provide insights into the type of monitoring systems which would be suited to the context of South Africa. The research has shown that a value-added monitoring system can be valid and reliable. However, there is room for improvement:

- The sample was restrictive not only in terms of size but also the demographic characteristics of schools which exist in South Africa. In this sample only urban and

peri-urban schools from one province were included. Urban, peri-urban and rural schools in other provinces were not included and this is seen as a limitation.

- The evaluation process, in which specialists in the field of education and psychology are consulted, could be extended to suggestions on what could be included in order to make inferences in terms of curriculum validity and content-related validity stronger. Here the selection of the specialists or the way in which the specialists are sampled will have to be done with care and with a specific purpose in mind.
- Follow-up interviews with National Department of Education officials could be undertaken in order to ensure that the specifications of the monitoring envisaged would comply with policies on a national-level. Furthermore, units directly involved in the implementation of the *Systemic Evaluation* and *Integrated Quality Management System* (which includes *Whole School Evaluation*) should also be included in order to add additional depth to the research.
- In addition to officials in the Provincial Department of Education it may have been beneficial to include officials working on the district-level as these officials would have more grass roots knowledge and potentially could provide valuable insights into how the envisaged monitoring system could be used in the varying schools contexts.
- The Rasch analyses could be extended to include equating items from different assessments and exploring the differential item functioning. Essentially, equating draws on item response theory where equating items from different grade assessments means that the items are linked. The difficulty of the items and the ability of the learners can then be put on the same scale (as was discussed in Chapter 5 and Chapter 7). By means of this analysis, potentially weak items can be identified and the ability of the learners ascertained. The results could then feed directly into topics for intervention programmes. The differential item functioning on the other hand, would detect bias in the items, specifically with regard to gender and cultural groups. The analysis of how items are performing for boys and girls and across racial groups would strengthen claims of cultural validity and would identify items which are not working well for the different groups. If these items were identified, changed or removed, then the assessment would be better in the long run and would result in an assessment which could be used in all contexts.
- The analysis of missing data. Although beyond the scope of this dissertation it would have been interesting to draw a distinction in the data between missing in terms of not reached and missing in terms of had an opportunity but did not answer. This could provide important information in terms of what learners can do and what they preferred not to do possibly due to time constraints but also due to their inability.

This research was essentially exploratory in nature and the sample sizes chosen were adequate for the nature of the research, but additional sampling for the qualitative component would have been beneficial in adding depth to the insights already gained. Challenges were also encountered in the analysis undertaken for the second main research question as was addressed in Chapter 8. The second main research question ***which factors could have an effect on learner performance and therefore inform the design of the monitoring system*** was essentially addressed by means of multilevel analysis. The following insights can be mentioned:

- Multilevel analysis should include as few variables as possible that explain the most variance especially on the school and classroom-level. However, a limited number of possible variables could be explored in this research, due to limited number at the upper levels. The limited number at the upper levels also made the investigation of random components impossible.
- The multilevel nature of school effects focuses on the interaction between the school and classroom-level. The possibility also exists to explore the link between the school and other levels such as the district and Provincial Education Departments.
- Only direct effects are taken into account, which leaves the researcher to hypothesise the indirect effects. Perhaps structural equation modelling would have been appropriate as an initial departure point for identifying factors on one level (in conjunction with correlation analysis) before including these variables in a multilevel model. Multilevel analysis is ideal when cross-level interactions and direct effects are of interest. However, the indirect effects of variables may provide valuable information which could inform the development of the monitoring system for the context of South Africa.

9.3.2 Reflection on this research in of light school effectiveness research

The use of school performance data has great potential to contribute to improvement efforts in education but at the same time, if handled ineptly, the research could prove to be irrelevant or create a situation which would have been better avoided altogether (Wyatt, 1996). Nevertheless, it could be said that the quality of learning is determined by the quality of education provided by schools, especially what learners do in the classroom. Teaching and learning should be an interactive process. Schools need to know how their learners are progressing and the difficulties that are experienced with regard to learning so that the needs of the learners can be met (Black & Wiliam, 1998). Adequate monitoring systems could be of use in this regard.

Monitoring in education is important, as it is a way in which to formally regulate levels of quality in education, it provides a mechanism to hold stakeholders accountable and it provides the impetus for ongoing improvement in education (Scheerens et al., 2003). Furthermore, given the amount of financial and human resources, which is put into education, the effects of education should be considered (Sammons, 2006). If it is said that education should prepare learners for the world of work and if the resources are allocated to this, then this implies a holistic view in terms of personal development, citizenship and indeed the necessary skills needed to succeed in the labour market (Luyten et al., 2005). Thus the resources allocated by government would show some return. If this is the case, as is in many countries around the world, then assessing the extent to which these goals are met is essential. The use of and distribution of resources are also linked to the relevance of educational objectives and whether these objectives are in reality attained. Also, the fair distribution of resources, especially in South Africa, is paramount likewise how these resources are translated into economic benefit (Scheerens et al., 2003).

The main aim of using school effectiveness research as a departure point was to contribute to the discourse of school effectiveness in a developing world context as opposed to a developed world context. Different types of monitoring systems have been discussed, namely the ZEBO-project in the Netherlands, the VCE data project in Australia, and the ABC+ model in the United States (Chapter 2). The need for projects such as these arose out of policy initiatives undertaken by local and national governments. The aim of these projects was to develop tools which schools could use for self-evaluation purposes so that adequate interventions could be put in place if need be. These projects were all initiated in the developed world context and indeed provide valuable information on how a monitoring system based on sound research should be approached. However, according to Doran and Lockwood (2006, p. 205) school effectiveness “decisions have hinged upon levels and changes over time in aggregated achievement measures for successive cohorts of different students and ranks of schools based on these measures.”

School effectiveness research has been criticised in the past and the use of value-added monitoring systems has been suggested to counter some of the critisms - MidYIS is an example of this. By means of making use of value-added results, fair comparisons can be made as low ability learners are compared with low ability learners (CEM, 2002m). The way in which value-added measures are used in order to produce the necessary information is of vital importance in order to find measures, which would best suit the South African context. Moreover, different approaches can be applied in order to develop a system that is focused on the improvement of learners and quality of education by raising expectations regardless of

background characteristics. Two approaches have been discussed in Chapter 2, namely a curriculum-based approach and a developed abilities approach. Both approaches yield important information. The curriculum-based approach makes use of assessments that are grounded in the curriculum and are administered on an annual basis so that progress from one grade to another can be ascertained. A developed abilities assessment on the other hand can provide baseline information of skills, which the learners have already developed. These skills then form part of the cross-curricular skills that can be used to predict future performance. Regardless of which approach is preferable, measurement error, and low reliability may produce findings that are biased (Luyten et al., 2005). Thus it was of importance to base inferences about the sustainability of the monitoring system on sound psychometric theory.

One of the main concerns and indeed the motivation to undertake this research was the issue of quality education. If inferences are made by practitioners and policy-makers about quality education, some form of monitoring is needed. In the context of South Africa, a situation arises where secondary schools need information on the basic skills learners have upon entry into secondary school. These skills can be built upon, whilst problematic areas should be identified, and strategies developed to focus on identified areas. The lack of performance in international content-based or curriculum-based assessments as well as national content-based or curriculum-based assessments is a case in point.

South Africa has not performed well in international comparative assessments like the TIMSS studies in 2003, 1999 and 1995 (HSRC, 2006, Howie 1997, 2001) as well as the SACMEQ study (Moloi & Strauss, 2005) where South African learners performed well below the international averages and below those of many countries. Likewise, the South African learners performed well below expectation in the Systemic Evaluation in Grades 3 and 6. These results although disappointing could provide valuable insights, especially if the assessments used have a high degree of curriculum validity. The results could be due to learners being ill prepared in terms of the content areas in addition to being unable to achieve the expected assessment standards (National Department of Education, 2005b).

Monitoring systems do provide a vehicle in which key concerns arising from poor performance can be addressed by means of intervention programmes based on feedback. Furthermore, by means of making use of a developed abilities assessment, which has shown, to have curriculum relevance, basic skills in key learning areas can be assessed. It is accepted that “being effective is not the same as staying effective” (Luyten et al, 2005, p. 264) but in a country like South Africa an important starting point would be to draw on school

effectiveness literature in order to identify possible variables which are vital if the monitoring system is to work.

9.3.3 Contribution to scientific and practical knowledge

South Africa is a developing world, although it has been described as a curious mix of developed and developing worlds. However, the challenges facing South Africa are very similar to our African counterparts, especially in terms of education. Equity and redress has been the driving force behind educational reforms. The education reforms themselves have at times been met with extreme opposition on grass roots level. The change in curriculum from a content-driven curriculum to an outcomes-based curriculum is a case in point. Monitoring the quality of education is another point of contention, as was discussed in Chapter 1. However, monitoring of a system is an important means of assessing the effectiveness or health of a system. This is always a means to the end of identifying components that need to be improved upon.

According to Frederiksen and Collins (1989, p. 27):

> There are enormous stakes placed on students' performance on educational tests. And there are consequently enormous pressures on school districts, school administrators, teachers and students to improve on tests.

If it is said that there are pressures on the stakeholders to improve test scores then it makes sense that the system in which the stakeholders play a part will adjust curriculum and instructional practices to maximize the scores achieved (Frederiksen & Collins, 1989). If this is true, then a valid and reliable monitoring system can do much to assist in the process of effecting the necessary changes.

Porter (1991) is of the opinion that there are three reasons why a system of indicators would be used to evaluate school processes:

- Purely a descriptive function to direct school policy;
- To serve as an evaluative instrument which will perform a monitoring function;
- To provide explanatory information when goals are not reached.

The aim of this research was to explore monitoring systems based on sound indicators that would serve as an evaluative instrument so that schools would have the necessary information to effect changes. The aim was to explore a system which would provide

systemic and comprehensive information (Porter, 1991) and would be flexible enough to adapt to the context of the school in which it would be implemented (Bryk & Hermanson, 1993). The idea was to explore a system which would provide accurate, valid and reliable information to schools, a system in which the "pulse of academic outcomes" (Bryk & Hermanson, 1993, p. 460) as well as "key inputs and processes" (Bryk & Hermanson, 1993, p. 460) are monitored.

The suggestion has been put forward that a suite of instruments under the banner South African Secondary School Information System (SASSIS) should be developed. This would ensure that the monitoring system rooted in sound indicators "can be used instrumentally at any level (Bryk & Hermanson, 1993, p. 460). Although this research only focused on the learner, classroom and school-level this could be extended to the district, province and national-levels as well in that the data can be aggregated to be used at higher levels of the education system.

The aim and rationale of the monitoring system is that the quality of education has to be monitored in order to identify areas of strength as well as areas which could be strengthened. If this monitoring system is to be successful then the CEA, schools and education officials should form a collaborative partnership. If schools and education officials are to be empowered then they have to feel that they are an important part of the process. Although this research did not explore the use of intervention programmes, what does come out is the issue of how the schools are going to use the information which is provided by the system. Are the schools in a position to design and develop intervention programmes which will assist learners to grow academically? Furthermore, what role will the CEA or any other agency involved in serving schools and districts with this type of monitoring play in providing the information and facilitation of the development of intervention programmes based on the information received? Clearly, if the intervention programmes are to be implemented then the schools and education officials have to be part of the development process. Otherwise this becomes similar to many Government initiatives which are prescriptive rather then participative.

An important aspect is that although achievement is an important component of the monitoring system it is not the only component. Various other components are also important, such as learner motivation to achieve and to study further (as was seen in Chapter 2). Thus attitudinal information should also be collected, analysed and feedback given. In school effectiveness research it is accepted that non-cognitive variables can be just as important as cognitive variables (Luyten, et al., 2005; Van Damme, Opdenakker, Van

Landeghen, De Fraine, Pustjens and Van de gaer, 2006). In this regard this research does contribute in providing the initial ground work to include non-cognitive variables (see Figure 9.2).

Another area in which this research contributes to the field is in terms of implementation. If the system is to work then there should be minimal effect on school and education officials' activities. Time is an important component. Schools have set yearly plans and goals which have to be met and education officials have their duties to attend to.

This research has also made explicit how indicators of effectiveness have been chosen and that the feedback given should result in positive action being taken. The monitoring system explored in this research has shown potential as functioning in a similar way in all contexts present in South Africa. Due to the disparities present in South Africa it had to be shown that the instruments can be used across contexts.
Furthermore, in this study a conceptual framework was developed which draws on the work of Scheerens (1990) and includes literature from the developing world (see Chapter 3 and Figure 9.2 in this Chapter). The main idea is that the road to school improvement can be built on school effectiveness research. Scheerens (1998) states that monitoring and feedback, based on school effectiveness factors, are of key importance for improvement purposes. Here the emphasis is on providing good quality information upon which self-evaluations on the learner, classroom and school-level can be based. Furthermore, poor performing schools would want to improve but even schools that are performing well should seek to find avenues of improvement. West (1998, p. 769) is of the opinion that no school, no matter how effective, should be "satisfied with its current provision - even the most successful of our schools could, indeed must, continually seek out ways to improve quality of outcomes and the experience of its students".

Very often, the educator is overlooked in models of school effectiveness only including two levels namely the school and the learner. Very few studies include the educator as an additional source of variation (Luyten et al., 2005). In this research, an attempt was made to construct a three-level model based on literature, including the classroom-level (see Chapter 3 and Chapter 8 as well as Figure 9.2). Not only is the classroom-level seen as an additional source of variation, it is accepted that there are indirect influences of the classroom-level on school-level factors via educator behaviour (Luyten et al., 2005).

Quantitative methods are used almost exclusively (Luyten et al., 2005) in school effectiveness research. In this research, an attempt was made to include both qualitative and

quantitative approaches. Although these approaches were linked to specific questions, they can be extended especially if the classroom-level is to be included. Here interviews and observations can be used to deepen arguments and add substance to recommendations. By purposefully mixing and/or combining qualitative and quantitative methods (Johnson & Christensen, 2004) stronger inferences can be drawn. Here the aim would be to further identify what educator characteristics and instructional practices are associated with effective schools or educators (Doran & Lockwood, 2006).

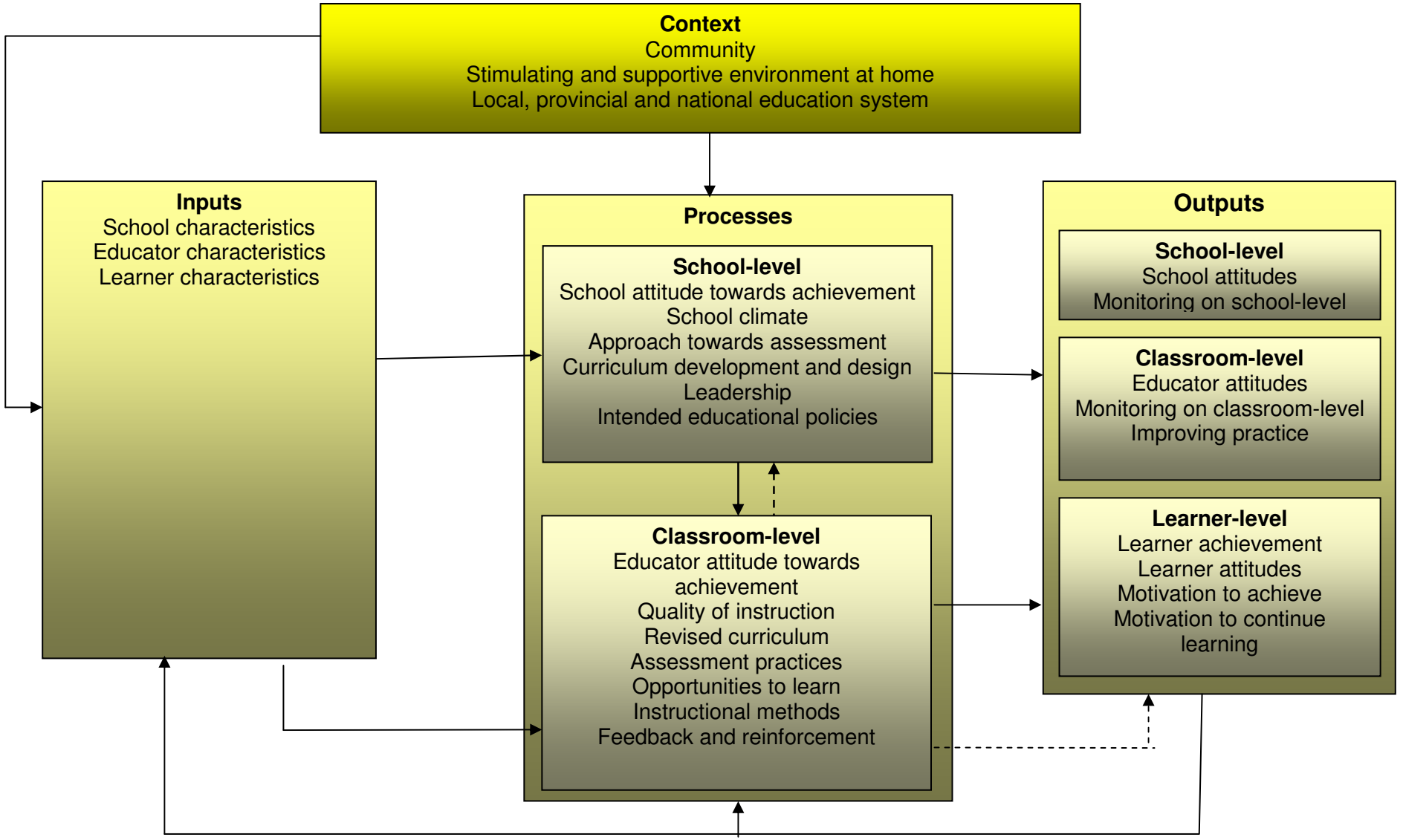

Figure 9.2 **Conceptual framework for monitoring education in South Africa (adapted from Scheerens, 1990)**

9.4 Recommendations

…evidence, like the truth, has many faces (Jansen, 2006, p. 35).

Jansen (2006) makes a compelling argument that evidence and indeed evidence-based research can be attainable by means of various methods. Policy should be informed by sound research practices but it should be acknowledged that the context in which decisions are made are not neutral and are often influenced by a number of factors. There will always be factors involved in making decisions. Recommendations related to monitoring systems are discussed in 9.4.1, while 9.4.2 elaborates on possible policy recommendations. The section concludes with recommendations regarding further research (9.4.3).

9.4.1 Recommendations and issues regarding monitoring systems

In this section only recommendations relating to the monitoring system will be put forward and discussed. Some of the recommendations put forward are not directly related to the findings in this research, but are extrapolations based on the research process.

Recommendation 1: Inclusion in monitoring systems of various levels

Monitoring involves the assessment of educational processes at various levels of the system. In this research only the school, classroom and learner-level were focused on. However, the aim of the assessment is to ascertain what needs to change and when it needs to change (Howie & Plomp, 2005). Throughout this research the idea of the nested nature of the education system has been emphasized. However, if the demands for quality education are to be met, then districts and provinces should also be included.

Recommendation 2: Design of a suite of instruments with a clear rationale linked to indicators

Monitoring the quality of education is a key focus point of this research, specifically ways in which the quality of education can be ascertained and observed over time. As the education system is a nested structure occurrences on one level of the system have an impact on the other levels. In order to design a comprehensive system, a suite of instruments (each targeting different levels and different sub-groups) is called for so that the data can provide valuable information and insights on each level. However, care must be taken that not too many instruments are developed resulting in an overly complex monitoring system, which would be difficult to manage.

Recommendation 3: Strong partnerships with and active collaboration by school leadership and educators

If the monitoring system is to be successful then the "buy in" of the various stakeholders, such as the CEA (or other agencies working with monitoring systems), school leadership and educators is of outmost importance. The stakeholders will need to know what the monitoring system is about and perhaps more importantly what is expected of them. If the monitoring system is essentially about school improvement then the schools will have to be empowered. A strong collaborative partnership between the CEA (and/or similar agencies) and will be needed.

Recommendation 4: Designing a system of reporting data that can be manipulated using aggregated and disaggregated data

The monitoring system should make provision for reporting data on various levels. On the classroom and school-level, individual learner performance may be of importance. However, it is plausible that on a provincial-level aggregated data is of more importance.

9.4.2 Recommendations and issues regarding policy

In the section to follow recommendations regarding policy will be addressed. The recommendations suggested flow from the research presented here but may not be directly related to the findings.

Recommendation 5: Systems should be identified which will assist schools in the process of self-evaluation

In South Africa the *Integrated Quality Management System* comprises *Systemic Evaluation* and *Whole School Evaluation* amongst others as was discussed in Chapter 1. As part of the Whole School Evaluation schools have to evaluate themselves on a yearly basis. Currently, in South Africa, monitoring systems aimed at assisting schools with the self-evaluation process do not exist. Thus it is recommended that the Government identifies systems which can be made available to schools. What is of importance is to be able to link self-evaluation data to performance data as this would provide schools with information as to how interventions, on a school and classroom-level, are impacting on learner performance. Furthermore, standardised forms of data collected on a school-level should complement the systemic evaluation process. A system such as the one under investigation in this research is such a possibility although further development and research is needed.

Recommendation 6: Government subsidies for development of valid and reliable systems to undertake self-evaluations

Currently in South Africa there is a lack of reliable indicators of quality education (Howie, 2002). The Systemic Evaluation component of IQMS is still in its infancy and is currently only available at primary school-level. For secondary education the Grade 12 exit examinations could be a possibility. It is recommended that Government subsidise the development of reliable monitoring systems at the lower secondary level as is the case in other countries such as the Netherlands. It is not possible to expect schools to design and implement monitoring systems as they simply do not have the capacity to do so. However, by means of Government subsidising the development of monitoring systems, schools will inevitably benefit.

Recommendation 7: Schools need to make policies regarding monitoring explicit

The type of monitoring that schools have in place could potentially be of benefit in terms of the self-evaluations that schools need to undertake. However, there is a lack of capacity at the school-level in terms of implementation. If Government is to institute the type of monitoring needed in order to adhere to the Whole School Evaluations then schools could possibly have a better idea of what is expected of them and what areas with the school system should be monitored. Furthermore, as long as the Whole School Evaluation process is linked to money which schools receive and not necessarily to an increase of the quality of education and facilitation of learner progress, this system will be problematic. The type of monitoring system envisaged is focused on determining the quality of education and the strengths as well as weaknesses of the school system.

9.4.3 Recommendations for further research

The section to follow includes recommendations for further research. Although the initial groundwork has been laid in this research, there are important considerations if the monitoring system explored is to be implemented on a wider scale.

Recommendation 8: Follow-up research activities should take place for the development of valid norms tables

If the SASSIS monitoring system is to be implemented across the country then research is needed for standardisation purposes, ensuring that it is working in the same way in the different provinces. Once the assessment is standardised then follow-up research should focus on developing valid norm tables which may serve as a guide to developing intervention programmes.

Recommendation 9: Interventions in line with taxonomies should be developed

By means of linking skills to a taxonomy such as the Anderson-Krathwohl taxonomy (also known as the revised Bloom taxonomy) targeted intervention programmes can be developed. This taxonomy includes cognitive processes as well as knowledge dimensions, which could serve as a guide for intervention programmes. Additionally this taxonomy allows for the inclusion of a third dimension, namely quality of the assessment (Killen, 2004).

Recommendation 10: Evidence-based intervention programmes should be explored

The design and development of the intervention programmes should draw heavily on the feedback that schools receive. Evidence suggests that feedback can be as harmful almost as often as it improves a situation. When designing and implementing feedback systems a number of cycles of evaluation and improvement may be needed. Under the right conditions feedback can have a substantial effect on the improvement of task performance (Coe, 2002). The aim of including feedback as a key area is to identify ways in which to maintain and improve the quality of schools. This aim arises out of the conviction that feedback is essential to learn in order to produce change (Coe & Visscher, 2002).

There are two additional recommendations. However, these pertain specifically to research design issues relevant to the current research.

Recommendation 11: The sample sizes at all levels should be increased

As this research was exploratory in nature the sample size (experts, schools, classes and learners) was appropriate. However, if inferences are to be made as to how the assessment is working across contexts then the full population of South African schools has to be included and the sample size as well as type of schools has to increase substantially. Furthermore, the sample sizes of the classroom and school-level have to substantially increase if reliable estimates of factors related to achievement are to be ascertained.

Recommendation 12: Existing national examinations should be incorporated as a data source

This research has shown that MidYIS, with adaptations, is valid and reliable for the South African context. However, in order to further elaborate on the predictive validity of the assessment, academic results attained from a common assessment are needed. It is possible to use the results of the Grade 9 and Grade 12 exit level examinations. In this research mathematics and English results were requested from the schools. The result was that the MidYIS assessment does explain some of the variation in the academic results. However, the amount of variation accounted for differed drastically between different

classifications of schools. By using a common assessment across schools, inferences based on predictive validity would be stronger.

9.5 The role of feedback and intervention

Value-added systems much like the systems used by CEM qualify as a performance feedback system in that:

> …a basic truism of learning implies that an individual student, not a student group, has increased in knowledge and skills during a specific period of time. As such, analytical methods concerned with student learning should reasonably reflect this basic principle and consider individual students as the unit of analysis with their growth trajectories employed as outcomes (Doran & Lockwood, 2006, p. 205).

The feedback used in these systems can be academic or non-academic or ideally both, which will assist schools to detect problems in functioning (Luyten, et al., 2005) If school performance feedback systems are to provide the information needed to assist schools then the research agenda could be guided by the following questions (Luyten et al., 2005):

1) How can the feedback be made accessible?
2) What information is deemed credible by schools?
3) What type of feedback is most accessible and easy to understand?
4) How can feedback systems be used to detect problems and find solutions?
5) What strategies for change are most effective?

Reflecting on the school as a system, perhaps the intervention based on feedback should also be conceptualized in terms of a hierarchical system. The ecology theory of human development elaborated on by Bronfenbrenner could be used. Bronfenbrenner's theory is comprehensive in nature and provides explanations of competence (Sontag, 1996). According to Bronfenbrenner (1975, p. 439) "an ecological perspective focuses attention on development as a function of interaction between the developing organism and the enduring environments or contexts". Furthermore, the ecological structure of the educational environment comprises various levels. If the intervention is to be effective then behavioural change should be viewed as being nested within a number of developmental contexts (Ramono, Tremblay, Boulerice & Swisher, 2005). How learners learn, according to Bronfenbrenner (1976), in educational settings is a result of two forces. The first is in the relationship between learners and their surroundings and the second includes the interconnections between the different environments. For Bronfenbrenner (1976, p. 5-6), the environment is a nested arrangement of structures comprising four levels:

i) The micro-system which is the immediate setting of the learner such as home or the classroom.
ii) The meso-system comprises the interrelations of the settings such as the school.
iii) The exo-system is an extension of the meso-system in which formal and informal social structures are included such as the community.
iv) The macro-system comprises overarching institutions of culture, typically the educational, legal and political system in which the micro-, meso- and exo-systems are the concrete manifestations.

If however, intervention strategies are designed from an ecological perspective then it may be beneficial to include as the first level the individual learner or a nano-level as Van den Akker (2003) calls it.

The interventions should engage the learner in a manner that will inspire skills which will be used in life and not merely skills as narrowly defined by the curriculum. This implies seeing learning "not simply as a high score on a test or assignment, but should involve increasing possibilities for action in the world" (Barb & Roth, 2006, p. 11). This is also in line with many education systems around the world where schooling is seen as a training ground for the world of work. The ecological view of learning is therefore useful in that it allows the developing of content that has "cross-textual value" (Barb & Roth, 2006, p. 3).

The conceptual model used in this research (Figure 9.3) can be viewed from the perspective of a monitoring system on a national-level. The indicators included in the model can be relevant for the various school contexts which exist in South Africa. How the schools respond to the data they receive is something different all together as some indicators may be more important than others. Thus a school-based model for improvement, reflecting the school context, can be developed which draws on the monitoring data received (see Figure 9.4).

Data are collected and processed by an external agency such as the CEA. Based on the data feedback on key indicators is given. This would create some pressure for the school to try and improve performance on the indicators. This pressure could culminate in the development of intervention programmes targeting key indicators on which schools (or a particular school) need to improve upon. The intervention programmes are developed and implemented over time. Data can be then collected again and processed by an external agency resulting in and feedback to the school(s).

Thus the school can monitor and evaluate whether the intervention programmes initiated made a difference i.e. undertaking a self-evaluation. For this purpose an evaluation model can be developed (see Figure 9.5). The evaluation model includes intended inputs, processes and outputs in addition to actual inputs, processes and outputs. The school would then be in a position to assess whether what was planned materialised. Furthermore, this would provide additional information on how intervention programmes can be improved upon for future use.

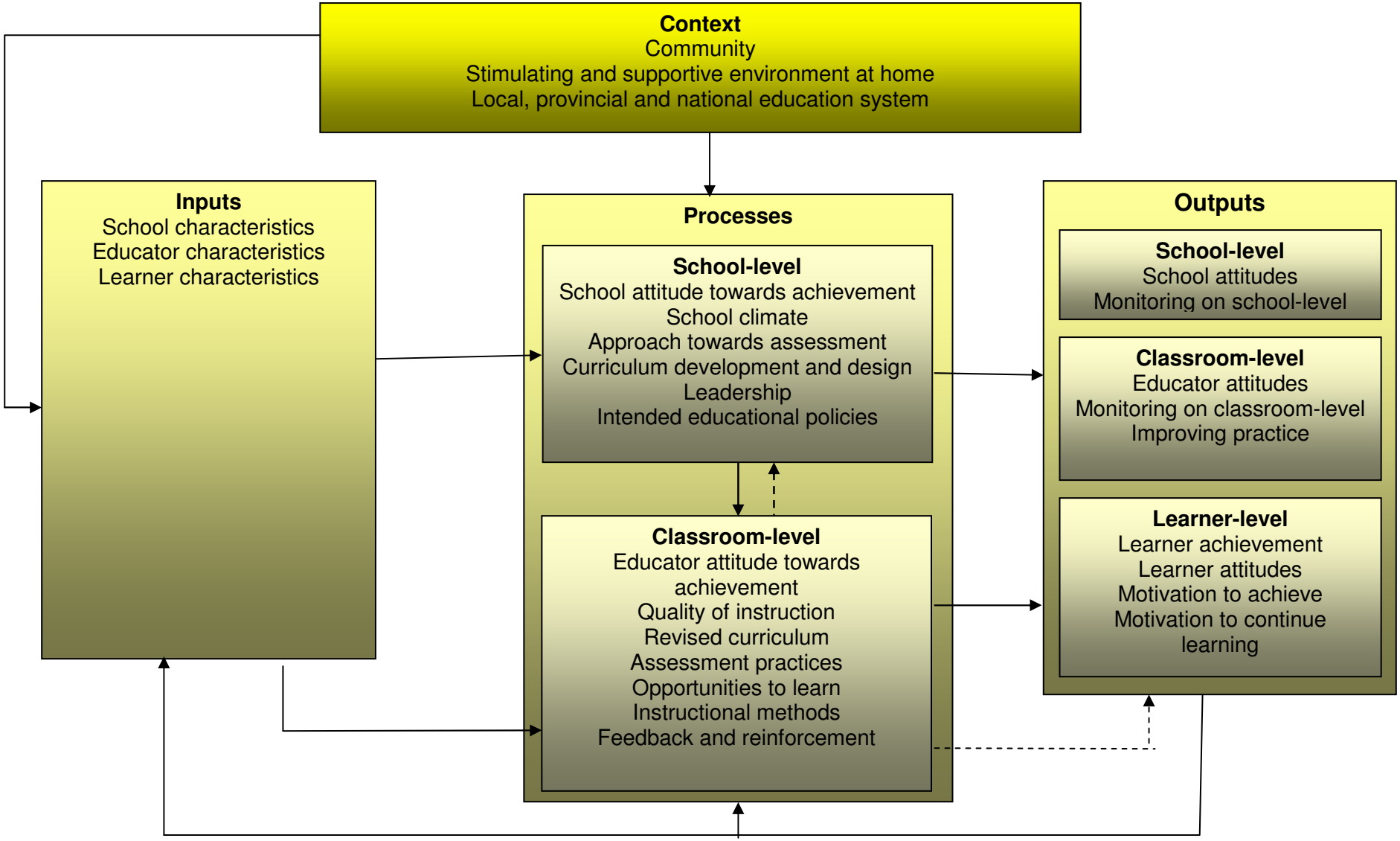

Figure 9.3 **Conceptual framework for this study on monitoring the quality of education**

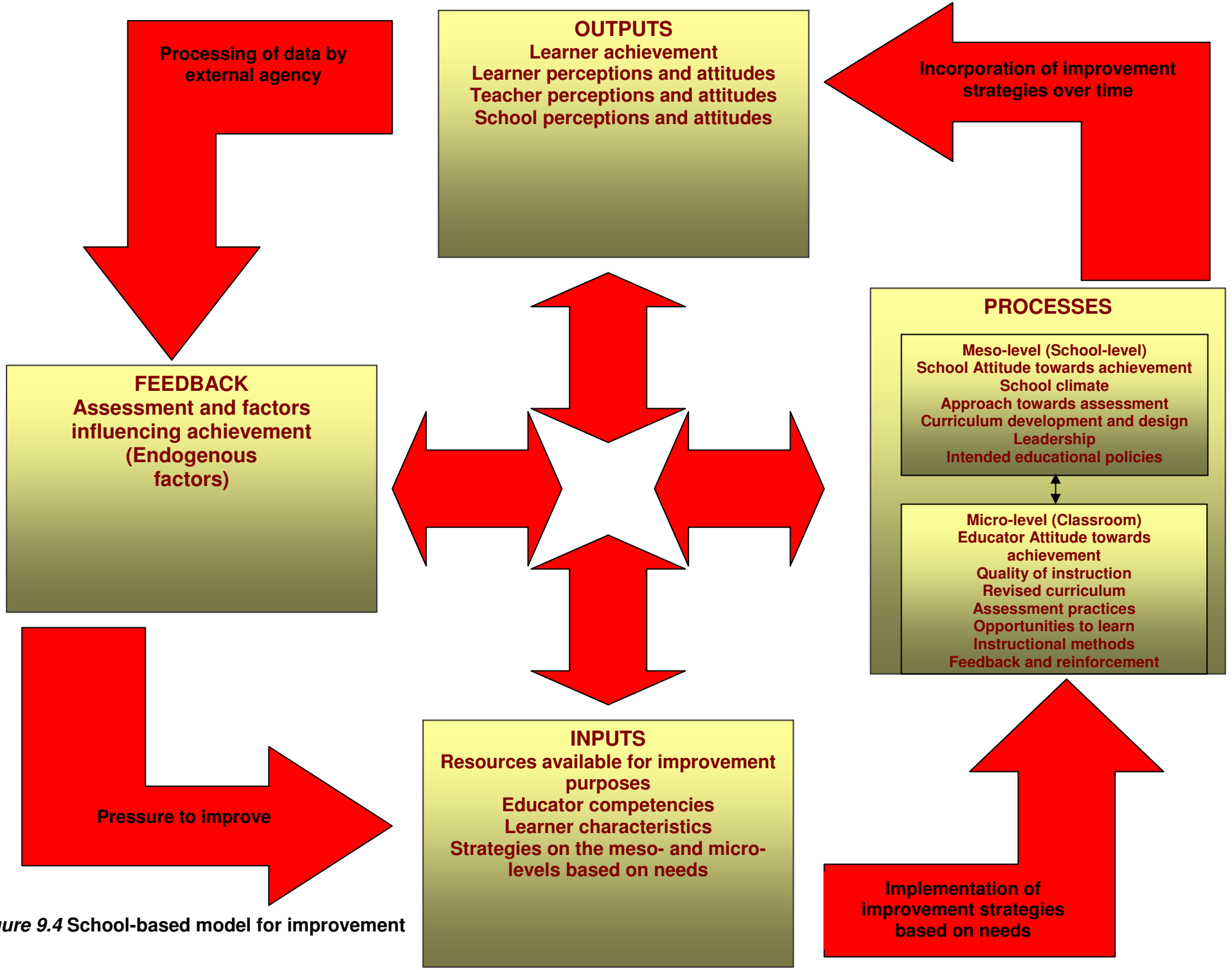

Figure 9.4 School-based model for improvement

Figure 9.5 taken from the countenance model of Stake (1968) illustrates the relationship between what would be originally intended with the intervention programmes and what actually may happened when the programmes are implemented.

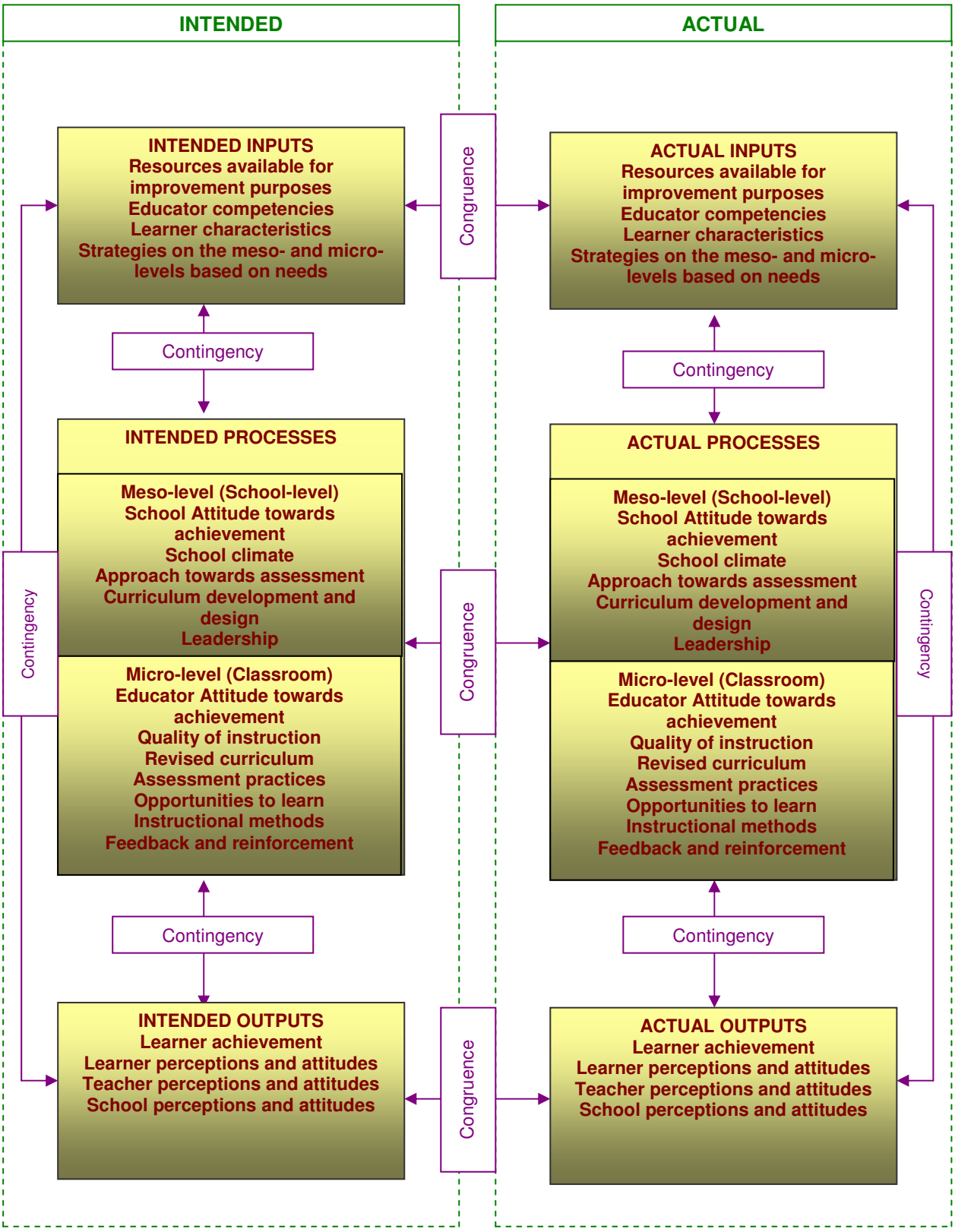

Figure 9.5 **Evaluation model for school-based improvement**

Stake (1968) uses antecedents to indicate what goes into the system or, as in Figure 9.5, inputs. He specifically refers to antecedents as background information. However this can be modified for the purpose of this research: what goes into the system goes through processes of teaching and learning and results in a number of outcomes. Instructional transactions form part of the processes, while outcomes specifically refer to what is achieved. Congruence in Figure 9.5 refers to whether what was intended actually occurred. Thus to be congruent the intended inputs, processes and outputs would have to come to pass (Stake, 1968). Contingency refers to the relationships among the variables characterised under inputs, processes and outcomes (Stufflebeam & Shinkfield, 1984).

Stake's thinking is important as the intervention programme will have intended outcomes which are based on assumptions about certain inputs and processes. There is a relationship between the inputs and processes and the processes and outcomes. These have to be identified and recognised if the intervention programme is to be a success. Ideally there would be an orderly cyclic process of developing effective education and, in the case of remediation, an effective intervention programme. Designing and developing an effective monitoring system or intervention programme is never achieved in one try. Rather, the development activity takes the form of a cyclic approach, in which development is undertaken. This is implemented followed by evaluation which results in revision and further development work. This is in essence the approach applied in design research. According to Van den Akker, Gravemeijer, McKenney and Nieveen (2006, p. 2):

> By carefully studying progressive approximations of ideal interventions in their target settings, researchers and practitioners construct increasingly workable and effective interventions, with improved articulation of principles that underpin their impact.

Design research aims to develop theories based on empirical evidence through the process of learning as well as the vehicle used to support the process of learning (Van den Akker et al., 2006). The process underlying design research can be characterized as follows (Edelson, 2006):

1) It is research driven and thus draws on prior research.
2) The research process is systematically documented.
3) The design is developed based on research and knowledge of the context is implemented.
4) The implementation is followed by formative evaluation in order to identify any weaknesses in problem analysis, design solutions or design procedure.

5) There are iterative cycles of design, implementation and evaluation.
6) Generalisations can be made in the form of theories, design frameworks or design methodologies.

9.6 Conclusion

> *The shortest distance between two points is still under construction – Noelie Altito* (Genn, 2007).

The aim of this PhD research was to explore the feasibility of using a value-added monitoring system for education developed in the United Kingdom. As the research continued the possibilities of what such a system could mean for South Africa presented themselves. Thus this was the first step in what appears to be the beginning of a very exciting journey.

A national monitoring system is proposed, a system which is not managed at the national-level, but possibly subsidised by the Department of Education to meet the goals of education that the government has identified. However, if the monitoring system is to work then the correct foundation has to be provided. In order to have, in the end, a system of high standards and quality, the following should be carefully considered and reported (Posthethwaite, 2004):

- The aims of the system should be explicated stated and ideally should be relevant to theory and policy. The aims of the system should be operationalised into good research questions;
- Descriptions of the target population should be elaborated on in terms of defined population, desired population and results for exclusions given;
- The sample should be specified as well as methods of sampling employed. Sampling weights should be used in order to correct for disproportionality among sampling strata. Sampling error and response rates should be reported;
- Translations of instruments have to be verified and the process adequately described;
- The assessments should be appropriate and domains clearly defined. The validity and reliability strategies should be reported;
- If questionnaires are included, the items should adequately cover the research questions and the variables be defined. The questionnaires should be piloted as an additional stage of refining;

- How the data were collected is of utmost importance and descriptions should be given in terms of the manuals used, tracking forms used, missing data and quality control mechanisms;
- The quality of the data should be described, i.e. the data entry program should be described and consistency checks elaborated on;
- The analysis techniques used should be appropriate and standard errors reported;
- The reports written should be clear and the relevant issues should be adequately addressed.

In conclusion

> As the number of school aged children has grown rapidly world-wide and the demand for the provision of both primary and secondary school has increases at an even greater rate, it has gradually become essential to monitor educational standards.
> (Keeves, Lietz, Gregory & Darmawan, 2006, p. 110)

In this context, the research presented in this book is just the first stage of a long route South African education has to travel in order to reach world-class quality.