Factors Influencing Pupil Achievement in SACMEQ II - Botswana: An Application of Structural Equation Modelling

Kim Cuc Nguyen

Margaret Wu

Shelley Gillis

The University of Melbourne, Australia

Paper presented to the International Invitational Educational Policy Research Conference, Paris, France, 28 September – 2 October, 2005.

Abstract

This paper examines the extent to which factors related to pupil, home, teacher and school background influence pupil achievement in mathematics and reading literacy in Botswana, one of the fourteen countries represented in the SACMEQ II investigation. The process of developing and examining a hypothesized causal model for explaining pupil achievement in Botswana was presented in detail to enable future researchers to replicate the structural equation modelling (SEM) procedures for each of the remaining thirteen countries. School location, family socio-economic status (SES), total school resources and class repetition were found to be the most important factors associated with pupil achievement in Botswana. The direct and indirect effects of the background characteristics on pupil achievement were explored and examined. As the analyses enabled causal effects to be identified, a series of specific intervention strategies were recommended for priority implementation in order to enhance future pupil achievement in Botswana.

Introduction

In collaboration with 15 Ministries of Education, the International Institute for Educational Planning (IIEP) formed the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) to collect achievement data among pupils within its member countries. A major driving force behind the formation of the SACMEQ was to assist educational planners with monitoring and evaluating the quality of basic education within the participating countries. Through the collection of data on pupil achievement levels, as well as the contextual and background factors from pupils and schools among the 14 participating countries, it was thought that decision makers could use the results to plan and improve the quality of education within the participating countries. Although the SACMEQ studies were not designed to explicitly explore the factors influencing pupil performance, the two data sets released by the IIEP contained a number of background variables and scales which could be used to address the issue of why pupils with a particular set of background characteristics, in a particular country, performed at a particular level. An understanding of factors that influence pupil achievement and the relative effects of each factor can inform education policies (Rothman and McMillan, 2003).

However, it was noted by researchers that "the main challenge of studies on schooling is to make some attempt to disentangle the overlapping effects of the home and school on educational outcomes" (Murimba, Moyo, Pfukani, Machingaidze and Mtembo 1994; Kulpoo, 1998). To pursue this vital question and take the "challenge" of disentangling the overlapping

effects, this study examined factors that influenced pupil achievement in reading and mathematics in Botswana, one of the fourteen countries represented within the SACMEQ II investigation. In particular, it examined the extent to which factors related to pupil, home, teacher and school background influenced pupil achievement. The explicit questions addressed in this study were:

- To what extent do pupil, home characteristics, teacher and school related factors influence end of primary school pupil achievement in mathematics and reading literacy in Botswana?

- How do these factors interact with each other to influence pupil achievement (the latent variable combining mathematics and reading)? What are the most important factors that explain pupil achievement levels?

Factors Influencing Pupil Achievement

This section briefly summarizes an overview of research on factors influencing pupil achievement. Research conducted by SACMEQ countries and developing countries was of particular interest. The findings presented in Keitheile and Mokubung's (2005) report on Botswana were also reviewed. Past researchers have often grouped the variables thought to influence pupil achievement in reading and mathematics into pupil, home, school and teacher related factors (Falayajo, Makoju, Okebukola, Onugha, and Olubodun, 1997; Keeves, 1972; World Bank, 2004). Therefore the factors reviewed are presented in accordance with each of these four groups of factors.

Pupil Related Factors

Gender

Gender has been one of the most frequently reported factors when examining differences in pupil achievement (Nguyen and Fahey, 2001; Nguyen, 2002; Rosier, 1978; Williams, Long, Carpenter and Hayden, 1993; OECD, 2004; World Bank, 2004). However, past research has not yielded consistent results regarding gender differences in pupil achievement. Findings of gender effects on pupil achievement have varied according to school level (primary or secondary), subject (maths or languages), time and country. Among countries involved in SACMEQ I, findings on the relationship between gender and pupil achievement have varied across countries. Milner, Chimombo, Banda and Mchikoma (2001) reported that in Malawi, boys did better than girls in reading, while Kulpoo (1998) found that in Mauritius, girls tended to perform slightly better than boys in reading. A UNESCO report (1994), however, pointed out that there was no gender difference in reading in Zimbabwe (Murimba et al., 1994). Keitheile and Mokubung (2005) indicated that in Botswana girls outperformed boys in both maths and reading, and the difference between reading scores for boys and girls was significant.

Pupil Age

Keitheile and Mokubung (2005) reported that the average age of a standard 6 pupil was expected to be approximately 139 months. The mean age of standard 6 pupils in

Botswana was 157.8 months, suggesting that they were about one and a half years older than expected. Research conducted by the World Bank in Vietnam found that pupils who were older than their class peers tended to have lower achievement in both maths and language (World Bank, 2004). Research has also found that pupils in rural areas and pupils from low socio-economic backgrounds tend to be older than their peers in urban areas and from higher socio-economic backgrounds (Milner et al., 2001; World Bank, 2004).

Pupil Absenteeism

Nguyen (1998, 2002) found that pupils in rural areas tended to be absent from school when it was production season, and this was one of the causes of low achievement and dropping out of school. A similar hypothesis was raised in the reports of Keitheile and Mokubung (2005), Milner et al. (2001), and Murimba et al. (1994). In Botswana, although the rate of absenteeism was two fifths of a day in the month preceding data collection in 2000 and was considered very low, absenteeism varied across regions.

Repetition

Repetition has also been found to be a serious problem among SACMEQ countries (Murimba et al., 1994; Milner et al., 2001). The repetition rate in Botswana was considered to be relatively high. Thirty one per cent of pupils reported that they repeated class at least once. Repetition rates varied across regions (Keitheile and Mokubung, 2005). Past research has produced consistent correlations between repetition and pupil achievement (World Bank,

2004; Tran, 1994). In addition, research findings have indicated that pupils from rural or remote areas and from families with low socio-economic status are more likely to repeat (World Bank, 2004).

Family Related Factors

Family Socio-economic Status (SES)

Family socio-economic status is one of the most frequently examined factors in studies exploring differences in pupil achievement, and has been found to be the most important factor associated with pupil achievement among both developed and developing countries. To disentangle the overlapping effects of SES and other factors, much past research has partialled out SES. In many cases the effects of other factors decreased after controlling for the effect of SES, suggesting that the effect of SES on pupil achievement was mediated through other variables. It has also been found that families in urban areas tend to have higher possession index than those in rural areas (Milner et al., 2001; World Bank, 2004; Keitheile and Mokubung, 2005). This suggests that the effect of location on pupil achievement might be mediated through SES.

Regularity of Meals

The average meal index for pupils in Botswana was 10.7, which was considered relatively high. It should be noted that those pupils who ate no meals at all for the whole week were given a score of 3, whilst those who ate all meals were assigned a score of 12. While

there was little variation in the average number of meals across regions, it was clear that the mean index for the West region was lower than the average meal index for Botswana (Keitheile and Mokubung, 2005). A number of studies have found a relationship between low nutrition or lack of food and low pupil achievement (Griffin, Woods, Dulhunty, Nguyen and Calvitto, 2003; World Bank, 2004). People in rural and remote areas and people with low SES have also been found to be more likely to suffer from hunger than those from urban areas and those with greater wealth (Nguyen, 1998, 2002; World Bank, 2004).

Home Interest

In Botswana the extent to which parents assisted their children with homework also varied across regions. Gaborone, an urban region, had the highest percentage of pupils who received assistance from their parents (Keitheile and Mokubung, 2005). Family attention to pupil achievement on pupil achievement was examined in the World Bank project monitoring student achievement in Vietnam (World Bank, 2004), and by Monitoring of Learning Achievement (Falayajo et al., 1997). Family attention was found to be slightly correlated with pupil achievement in Vietnam (World Bank, 2004) but not in Nigeria (Falayajo et al., 1997). Home interest is a factor that can be manipulated, therefore it is important to explore its relationship with pupil achievement.

Speaking English Outside School

In countries where English is the language of instruction in school, it has been repeatedly found that speaking English at home influences pupil achievement (Kulpoo, 1998; Rothman, 2002; OECD, 2004)). In Botswana, English was spoken least in the Central North region (Keitheile and Mokubung, 2005). People who did not speak English were more likely to be from a low socio-economic background (Rothman, 2002; Rothman and McMillan, 2003; OECD, 2004)

Teacher Related Factors

Teacher Qualifications, Teacher Experience and Teacher Test Scores

In Botswana, Standard 6 teachers had an average of 10.8 years of teaching experience and 2.2 years of teaching training. The teachers' scores in both reading and mathematics were much higher than those of their pupils, with national means for reading and mathematics being 757.7 and 753.3 respectively. There was considerable variability in teacher maths and reading scores across regions. Teachers from Gaborone, an urban area, tended to perform better than teachers in the Central South and the West (Keitheile and Mokubung, 2005).

The relationships between pupil achievement and teacher characteristics such as qualifications, experience and test scores in maths and reading have been well documented (World Bank, 2004; Falayajo et al., 1997). Research has indicated that teachers in urban areas outperform those in rural and remote areas in both reading and maths (World Bank, 2004).

Teacher In-service Training Courses

On average, teachers in Botswana had attended 2.7 in-service courses over a period of three years. Teachers in the Central North reported the lowest frequency of in-service course attendance. Since teachers' subject knowledge and qualifications have been found to influence pupil achievement (World Bank, 2004; Falayajo et al., 1997), it was hypothesized that the number of in-service training courses a teacher attended might influence pupil achievement.

School Related Factors

Pupil-Teacher Ratio

Pupil-teacher ratio is the total number of pupils in the school divided by the number of full–time equivalent teachers posted to that school. Therefore this factor could be considered as an indicator of school wealth in terms of school human resources. Griffin (2000) reviewed the data from 37 developing countries and found that a high pupil-staff ratio was one of the school factors that was associated with poor learning outcomes.

Class and School Resources

Classroom and school resources have been found to influence primary school pupil achievement (World Bank, 2004; OECD, 2004). Therefore it was also hypothesised that these factors might influence pupil achievement in Botswana.

School Location

School location has been one of the most studied factors associated with pupil achievement. Research has consistently found that pupils attending schools in urban areas perform at a much higher level than pupils attending school in rural areas (Keitheile and Mokubung, 2005; Milner et al., 2001; Kulpoo, 1998; Murimba et al., 1994; World Bank, 2004; Falayajo et al., 1997). A number of studies have also reported that the effect of school location on pupil achievement was indirect and was mediated by other variables such as family socio-economic background (Kulpoo, 1998).

In summary, the extent to which each of these factors has been studied and has been shown to contribute to pupil achievement is varied and often inconsistent across studies, however there is evidence that these factors are associated with pupil achievement. It was deemed worthwhile to explore if this were the case in Botswana. However, past research has also found evidence of interactions between factors. Therefore, the possibility that the effect of location on pupil achievement might be mediated through family related factors and pupil and other school related factors was investigated. Similarly, the possibility that the effect of family socio-economic status on pupil achievement might be mediated through other pupil, teacher and school related factors were also considered.

The Methodology

The Data

The data were taken from the archive maintained by the SACMEQ in association with IIEP. The data files contained weighting information to allow for complex sampling designs and to adjust statistical significance and standard errors accordingly. Information about the data and requests for access were made from the SACMEQ website http://www.sacmeq.org/access.htm. This paper reports on the findings from a secondary data analysis for the SACMEQ II study only (refer to Ross, Saito, Dolata, Ikeda and Zuze (2004) for an overview of the dataset).

The Variables

Factors that might influence pupil achievement in Botswana were selected on the basis of the literature reviewed. A description of each of the variables that examined, as well as the type of variable, the name used in the analysis, the data file code and a description of the data is presented in *Table 1*.

Table 1 about here

The Analysis

The descriptive statistics for the dependent (pupil achievement in maths and reading) and independent (those related to pupil, home, teacher and school background characteristics) variables for Botswana have previously been reported by Keitheile and Mokubung (2005). The mean scores for reading and maths in Botswana were 521.1 and 512.9 respectively. That is, pupils within Botswana performed above the SACMEQ II average in both reading and mathematics. Overall, the descriptive statistics for Botswana revealed sufficient systematic variation justify further bivariate and multivariate analyses.

To answer the research questions, two types of analysis were conducted: bivariate and multivariate. Each of these analyses is described below.

Initially, the relationship between each background variable and pupil achievement levels in *Table 1* was examined using a series of correlational analyses. Correlation is primarily concerned with describing the strength and direction of a relationship between two variables. The strength and direction of the relationship can be expressed by means of a correlation coefficient (Tilley, 1994:109). The closer the correlation coefficient is to -1 or +1, that is the further away from 0, the stronger the relationship. As a result of these initial exploratory analyses, the variables that highly associated with pupil achievement were identified.

The results of these analyses were then used to develop a hypothesised model of causal relationships among the pupil, family, teaching and school background variables. The model was then tested using linear structural equation modelling (SEM) to assist policy makers in identifying appropriate intervention strategies to improve pupil performance in Botswana. This approach to assisting policy makers with decision-making was consistent with the recommendation by Postlethewaite (1997:7), in which he argued "causal questions are usually the most important to educational planners".

Structural equation modelling (SEM) employs a set of equations to formulate structural relations that are hypothesised to exist between a network of observed and latent variables (Bollen, 1989). It can be used to test causal theories by simultaneously applying multiple regression techniques and factor analysis. Compared to other analyses, SEM has a range of advantages as it goes beyond descriptive, univariate, and regression-type analyses (Bollen, 1989; Pedhazur and Schmelkin, 1991). It can simultaneously invoke measurement and regression procedures, thus enabling a researcher to isolate and analyse measurement error, unexplained variance and true variance at the same time (Smith, 2000). Furthermore, it can separate the effects of one variable onto another into direct and indirect effects.

The SACMEQ II data set had several weights. According to Ross et al. (2004:11), pupil weight was "calculated for each country from the inverse of the probability of pupil selection – and then adjusted so that the sum of the weights was equal to the number of pupils in the achieved sample." Prior to conducting secondary data analyses on the Botswana data,

the sample was weighted using the PWEIGHT2 variable as recommended by Ross et al. (2004). To ensure the pupil weight was taken into account it was necessary to calculate a covariance matrix using SPSS first and then calculate the direct and indirect effects in AMOS using the covariance matrix that was calculated using the PWEIGHT2.

The Results

The analysis is presented in two broad sections: (a) the bivariate analysis; and b) the multivariate analysis. In the bivariate analysis, the relationships of the background factors and pupil achievement in both maths and reading are reported and, in the multivariate analysis, a detailed evaluation of the model is presented, as well as policy recommendations for improving pupil performance in Botswana.

Bivariate Relationships between Background Variables and Pupil Achievement in Botswana

This section reports on the process of identifying factors associated with pupil achievement in Botswana and hypothesising the interrelationships among these variables and pupil achievement. To identify the factors associated with pupil achievement in Botswana, Pearson's product moment correlations were computed. The results are reported in *Table 2*.

Table 2 about here

It can be seen from *Table 2* that pupil achievement scores in maths and reading were significantly correlated with pupil gender, pupil age, family socio-economic background, speaking English outside school, regularity of meals, home interest, class repetition, reading teacher qualification, maths teacher qualification, teacher reading and maths score, maths and reading teacher years of teaching, pupil-teacher ratio and school total resources. The four variables reading teacher in-service courses, maths teacher inservice courses, reading total class resources and maths total class resources were found not to be associated with pupil reading and maths scores. The observation that class resources were not associated with pupil achievement can be explained by the finding reported by Keitheile and Mokubung (2005) that there was not much disparity across regions.

From *Table 2* it can be seen that correlations between a number of variables (e.g. reading teacher qualification, maths teacher qualification) and both maths and reading achievement levels were relatively small (less than 0.10). Because pairs of variables with relatively small correlations do not contribute greatly to explaining variance, it was decided that only variables with correlation coefficients larger than 0.15 with the dependent variables would be considered in the structural equation modelling. The list of variables for further examination using SEM is displayed in *Table 3.* For the convenience of writing the structural and measurement equations, variable numbers were also included. *Table 2* also shows correlations between maths and reading achievement scores. The correlation between maths and reading achievement was relatively high (0.77). It appeared as though these two variables

were underpinned by a latent variable, which has been named pupil academic achievement (ACHI).

Table 3 about here

Interrelations among the Background Variables and Pupil Achievement: An Examination of the Botswana Data

Hypothesised Model of Factors Influencing Pupil Achievement in Botswana: Application of SEM

The review of literature on influences on pupil achievement presented earlier not only demonstrated evidence of the influences of factors on pupil achievement but also suggested possible interactions between these factors. The following propositions were established:

1. School location, family socio-economic status, regularity of meals, speaking English outside school, pupil class repetition, pupil age, teacher reading score, school total resources, and pupil-teacher ratio influenced pupil achievement.

2. School location influenced family socio-economic status, regularity of meals, speaking English outside school, pupil class repetition, pupil age, teacher reading score, school total resources and pupil-teacher ratio

3. Family socio-economic status influenced regularity of meals, speaking English outside school, pupil class repetition, pupil age, teacher reading score, school total resources, and pupil-teacher ratio.

The propositions outlined suggested that the factors presented in *Table 3* might influence pupil achievement simultaneously in their comprehensive interactions. While there have been a number of studies that either tested or suggested the possible interactions between factors in influencing pupil achievement, none have addressed all these factors simultaneously in interaction. Thus, while there is evidence that each factor might have both direct and indirect influences on pupil achievement, the extent to which each factor directly and indirectly influences pupil achievement has not been examined adequately. That is, there is a need to test the model developed when linking all the propositions from 1 to 3 into one model. Linking all the propositions, a saturated hypothetical model of factors influencing pupil achievement in Botswana was developed and is presented in *Figure 1*.

Figure 1 about here

In *Figure 1* the 11 rectangles represent the 11 observed variables reported in *Table 3*. The one circle represents the underlying latent variable, pupil academic achievement (ACHI). A straight, one directional arrow indicates that the variation in the variable pointed to is explained by the variation of the variable that the arrow originates from.

From *Figure 1* it can be seen that:

- There were 11 observed variables involved in the model.

- There was one latent variable: pupil academic achievement (ACHI). The latent variable, pupil academic achievement (ACHI), had two observed variables: pupil reading score (ZRALOCP) and pupil maths score (ZMALOCP). There were no correlations allowed between the residuals of these observed variables.

- School location (ZSLOC) was the exogenous variable. That is, there were no assumed variables that influenced this variable.

- Family socio-economic status (ZPSES), regularity of meals (ZPREGME), speaking English outside school (ZPENGLIS), pupil class repetition (ZPREPEAT), pupil age (ZPAGEMON), teacher reading score (ZRALOCT), school total resources (ZSRTOT22), and pupil-teacher ratio (ZSPTRATI) were endogenous variables.

- The variable pupil academic achievement (ACHI) was a dependent variable.

The following relationships were proposed and were tested in the saturated model. It was proposed that:

- School location (ZSLOC) had a direct effect on family socio-economic status (ZPSES), regularity of meals (ZPREGME), speaking English outside school (ZPENGLIS), class repetition (ZPREPEAT), pupil age (ZPAGEMON), teacher reading score (ZRALOCT), school total resources (ZSRTOT22), pupil–teacher ratio (ZSPTRATI) and pupil academic achievement (ACHI).

- Family socio-economic status (ZPSES) was influenced by school location (ZSLOC). In turn, family socio-economic status (ZPSES) had a direct effect on regularity of meals (ZPREGME), speaking English outside school (ZPENGLIS), class repetition (ZPREPEAT), pupil age (ZPAGEMON), teacher reading score (ZRALOCT), school total resources (ZSRTOT22), pupil–teacher ratio (ZSPTRATI) and pupil academic achievement (ACHI).

- Class repetition (ZPREPEAT) was influenced by family socio-economic status (ZPSES) and school location (ZSLOC). In turn, class repetition (ZPREPEAT) had a direct effect on pupil academic achievement (ACHI).

- Speaking English outside school (ZPENGLIS) was influenced by family socio-economic status (ZPSES) and school location (ZSLOC). In turn, speaking English outside school (ZPENGLIS) had a direct effect on pupil academic achievement (ACHI).

- Pupil age (ZPAGEMON) was influenced by family socio-economic status (ZPSES) and school location (ZSLOC). In turn, pupil age (ZPAGEMON) had a direct effect on pupil academic achievement (ACHI).

- Regularity of meals (ZPREGME) was influenced by family socio-economic status (ZPSES) and school location (ZSLOC). In turn, regularity of meals (ZPREGME) had a direct effect on pupil academic achievement (ACHI).

- Teacher reading score (ZRALOCT) was influenced by school location (ZSLOC). In turn, teacher reading score (ZRALOCT) had a direct effect on pupil academic achievement (ACHI).

- School total resources (ZSRTOT22) was influenced by school location (ZSLOC) and family socio-economic status (ZPSES). In turn, school total resources (ZSRTOT22) had a direct effect on pupil academic achievement (ACHI).

- Pupil-teacher ratio (ZSPTRATI) was influenced by school location (ZSLOC). In turn, Pupil-teacher ratio (ZSPTRATI) had a direct effect on pupil academic achievement (ACHI).

Examination of the Measurement Model

The first step in examining the full SEM presented in *Figure 1* was to assess the validity and reliability of the observed variables to define a given latent variable. Initially, the extent to which the variance estimates of the observed variables, pupil reading score (ZRALOCP) and pupil maths score (ZMALOCP), were determined by the latent variable, pupil academic achievement (ACHI), was examined. It was reported previously that the correlation between these two variables was 0.77, hence indicating that these two variables may have been underpinned by a single factor, that is, pupil academic achievement (ACHI). A path diagram of the proposed relationship between the two observed variables (i.e. pupil reading score and pupil maths score) and the latent trait variable (i.e. pupil academic achievement) is presented in *Figure 1* on the right-hand side of the diagram.

Schumacker and Lomax (1996:81) stated that:

> *The relationships between the observed variables and the latent variable are described by factor loadings. The factor loadings provide us with information about the extent to which a given observed variable is able to measure the latent variable. They serve as a validity coefficient. Measurement error is defined as that portion of an observed variable that is measuring something other than what the latent variable is hypothesized to measure.*

As previously stated, in the saturated model there were two observed variables (pupil reading score (ZRALPOCP) and pupil maths score (ZMALOCP)) measuring the latent variable, pupil academic achievement (ACHI). The equations underpinning the measurement model for pupil academic achievement are depicted in *Equation 1:*

$$\text{ZRALOCP} = \lambda_{5.7}\ \text{ACHI} + e_5 \qquad (1)$$

$$\text{ZMALOCP} = \lambda_{6.7}\ \text{ACHI} + e_6$$

where $\lambda_{5.7}$ and $\lambda_{6.7}$ are the coefficients relating ZMALOCP and ZRALOCP to ACHI respectively. The coefficient $\lambda_{x.y}$ represents the construct loading of V_x on V_y. e_5 and e_6 are the errors of measurement associated with the observed variables ZMALOCP and ZRALOCP respectively, whilst e_7 refers to the error of measurement for the latent variable ACHI.

The AMOS software program (Arbuckle, 2004) was employed to conduct the SEM analysis. The standardised factor loadings and the proportion of variance of each observed variable that was accounted for by the latent construct, as well as their measurement error, are displayed in *Table 4*.

Table 4 about here

The two observed variables, pupil reading score (ZRALOCP) and pupil maths score (ZMALOCP), had high factor loadings on the latent variable pupil academic achievement

(ACHI). It can also be seen that the loading of the variable pupil reading score (ZRALOCP) was higher than that of the variable pupil maths score (ZMALOCP).

The proportion of the total variance accounted for by the latent variable can be computed by squaring each of the factor loadings, summing them, and then dividing the sum by the number of observed variables (Schumacker and Lomax, 1996). The calculation of the proportion of the total variance accounted for by the latent variable shows how well the observed variables measure the latent variable (Schumacker and Lomax, 1996). The proportion of the total variance accounted for by the latent variable achievement (ACHI) was: $h^2=\sum\lambda_{i.7}^2/2=75.5\%$ (where i=5,6).

This means that the latent variable, pupil academic achievement (ACHI), accounted for 75.5 per cent of the total variance of the two observed variables (i.e. ZRALOCP and ZMRALOCP). The unique (residual) factor variance or the proportion of the total variance not explained by the latent variable can be computed as the average of one minus these squared factor loadings (Schumacker and Lomax, 1996). Thus, the proportion of the total variance not explained by the latent variable Pupil Academic Achievement (ACHI) was

$1-h^2=\sum(1-\lambda_{i.7}^2)/2=25.5\%$ (where i=5,6). This means that 24.5 percent of the total variance of the two observed variables was due to unique or measurement error.

Examination of the Structural Model

The next step in testing the relationships between the variables specified in the model was to examine the pattern of causal structure linking the exogenous, endogenous and dependent variables. The relationships between exogenous, endogenous and dependent variables specified in *Figure 1* is depicted in *Equation 2:*

$$ZPSES=\beta_{1.0}\,ZSLOC + e_1;$$

$$ZPENGLIS=\beta_{2.0}\,ZSLOC + \beta_{2.1}ZPSES + e_2;$$

$$ZRALOCT=\beta_{11.0}\,ZSLOC + e_{11};$$

$$ZREPEAT=\beta_{3.0}\,ZSLOC + \beta_{3.1}ZPSES + e_3;$$

$$ZPAGEMON=\beta_{4.0}\,ZSLOC + \beta_{4.1}ZPSES + e_4;$$

$$ZPAREGME=\beta_{8.0}\,ZSLOC + \beta_{8.1}ZPSES + e_8;$$

$$ZSRTOT22=\beta_{9.0}\,ZSLOC + \beta_{9.1}ZPSES + e_9;$$

$$ZSPTRATI=\beta_{10.0}\,ZSLOC + \beta_{10.1}ZPSES + e_{10};$$

$$ACHI=\beta_{7.0}\,ZSLOC + \beta_{7.1}ZPSES + \beta_{7.11}ZRALOCT + \beta_{7.2}ZPENGLIS + \beta_{7.3}ZREPEAT + \beta_{7.4}\,ZPAGEMON + \beta_{7.8}\,ZPAREGME + \beta_{7.9}\,ZSRTOT22 + \beta_{7.10}\,ZSPTRATI + e_7. \quad (2)$$

$\beta_{i.j}$ were structure coefficients of the path from variable V_j to variable V_i as presented in *Figure 1*. For example, $\beta_{7.1}$ represents the structure coefficient of the effect of ZPSES (V_1) on ACHI (V_7). e_7 was the error associated with the latent variable ACHI (V_7) and $e_{1,}$ $e_{2,}$ $e_{3,}$ $e_{4,}$ $e_{8,}$ $e_{9,}$ e_{10} and e_{11} are errors associated with the variables ZPSES (V_1), ZPENGLIS (V_2),

ZREPEAT(V_3), ZPAGEMON (V_4), ZPAREGME (V_8), ZSRTOT22 (V_9), ZSPTRATI (V_{10}) and ZRALOCT (V_{11}).

It was through these equations that the structure coefficients were computed, which indicated the strength and direction of relationships between variables. Each equation also contained an error or disturbance term that indicated the portion of the dependent variable that was not explained or predicted by the independent variables in that equation.

Model Fit

Inspection of the comparative fit of the model to the data using the Bentler-Bonett Normed Fit Index (*NFI*=0.85), the Comparative Fit Index (*CFI*=0.85) and the Bollen Incremental Fit Index (IFI=0.85) indices suggested a reasonable model fit (Schumacker and Lomax (1996) recommended that values close to 0.9 reflect a good fit). The *NFI* value suggested that the model was 85 per cent better fitting than the null model (i.e. where there are no relationships between the variables composing the model). Furthermore, the value of the model fit index, the Root Mean Square Error of Approximation *(RMSEA*=0.10) was 0.10. This also indicated a reasonably good fit to the data (Steiger, 1990). The goodness of fit index (GFI) and adjusted goodness of fit index (AGFI) were 0.94 and 0.88 respectively, also indicating a good fit. As these criteria are based on differences between the observed and the model-implied correlation or covariance matrix (Schumacker and Lomax, 1996), a good fit implies that there is not a significant difference between the fitted covariance matrix (model) and the sample covariance matrix obtained from the observed data. As such, both the model

comparison fit indices and the model fit indices suggested that the hypothesised model had a reasonably good fit to the data.

Direct and Indirect Effects

Examination of the direct and indirect effects may suggest paths that are not statistically significant and, thus, should be removed from the hypothesised model. This would also improve the fit statistics. In *Table 5* the direct, indirect and total effects of each factor in the saturated model on the variable pupil academic achievement (ACHI) are reported.

Table 5 about here

According to Schumacker and Lomax (1996), the direct effect of one variable on the other variables is the effect identified when controlling for other factors, and it is measured by the structure coefficients (β) between two variables. However, a variable can influence another variable indirectly, through one or more variables via their paths. For example, it can be seen from *Table 5* that school location (ZSLOC) indirectly influenced pupil academic achievement (ACHI) through family socio-economic status (ZPSES), regularity of meals (ZPREGME), speaking English outside school (ZPENGLIS), class repetition (ZPREPEAT), pupil age (ZPAGEMON), teacher reading score (ZRALOCT), school total resources (ZSRTOT22) and pupil–teacher ratio (ZSPTRATI). This was defined as an indirect effect,

and was measured by the product of the structure coefficients involved (Bollen, 1989; Schumacker and Lomax, 1996). The total effect of an endogenous variable on another variable is the sum of any direct and/or indirect effect that connects them (Bollen, 1989; Schumacker and Lomax, 1996).

Estimation of direct, indirect and total effects was used to facilitate understanding of the ways in which factors interacted with each other and the extent to which each variable influenced pupil academic achievement (ACHI). These effects were used to unlock the mechanism by which each factor influenced pupil academic achievement. In particular, the total effect of a variable on a dependent (or criterion) variable indicated the importance of that variable in explaining the variance in the dependent variable; the direct and indirect effects revealed the way that a variable interacted with other variables to explain the variance in the dependent variable.

It was noted from *Table 5* that the paths from school location (ZSLOC) on class repetition (ZPREPEAT), regularity of meals (ZPREGME) and pupil academic achievement (ACHI) were relatively weak and not statistically significant. Consequently, as school location (ZSLOC) did not have any direct effects on these variables, these paths were removed from the model. Furthermore, to develop a simpler model to explain end of primary school pupil academic achievement in Botswana, all paths with regression weights less than 0.10 were also removed from the model. For example, while the data in *Table 5* suggested that older pupils were less likely to successful in maths and reading, the effect was relatively

small (β=-0.08). Hence this variable was removed from the model. As a result of the examination, a simpler model of factors influencing pupil achievement in Botswana was re-specified and re-examined using AMOS (Arbuckle, 2004).

The Re-specified Simple Model of Factors Influencing Pupil Achievement in Botswana

The re-specified simple model comprised the following observed variables: school location (ZSLOC), family socio-economic status (ZPSES), speaking English outside school (ZPENGLIS), regularity of meals (ZPREGME), pupil–teacher ratio (ZSPTRATI), reading teacher score (ZRALOCT), total school resources (ZSRTOT22), class repetition (ZPREPEAT), pupil maths score (ZMALOCP) and pupil reading score (ZMALOCP). There was also one latent variable referred to as pupil academic achievement (ACHI). The solution to the re-specified simple model is presented in *Table 6* and *Figure 2*. Compared to the saturated model presented in the previous section, the fit statistics of this model were better. All of the model comparison fit indices such as NFI, CFI, and IFI were 0.91, indicating a good fit to the data (Schumacker and Lomax, 1996). The model fit indexes (RMSEA = 0.08, GFI = 0.96 and AGFI = 0.93) also indicated a good fit to the data. That is, both the model fit indices and the model comparison fit indices suggested that the re-specified simpler model fitted the data. The model explained 30 per cent of variance in pupil academic achievement. The direct, indirect and total effects of variables in the model on each other are presented in *Table 6*.

Table 6 and Figure 2 about here

The effect of school location (ZSLOC) on pupil academic achievement (ACHI) was indirect via family socio-economic status (ZPSES), pupil-teacher ratio (ZSPTRATI) and total school resources (TSRTOT22). That is, on its own, school location in Botswana was not an important influence on pupil achievement, but when combined with family socio-economic status, pupil-teacher ratio and total school resources, school location was an important influence on pupil academic achievement

The total effect of family socio-economic status (ZPSES) on pupil academic achievement (ACHI) was relatively large. Family socio-economic status (ZPSES) had both direct and indirect effects on pupil academic achievement (ACHI). The indirect effect of family socio-economic status (ZPSES) on pupil academic achievement (ACHI) was mediated through class repetition (ZPREPEAT), speaking English outside school (ZPENGLIS), regularity of meals (ZPREGME), total school resources (ZSRTOT22), and pupil-teacher ratio (ZSPTRATI).

The following variables had only direct effects on pupil achievement:

- Class repetition
- Speaking English outside school

- Regularity of meals
- Total school resources
- Pupil-teacher ratio
- Teacher reading score

The proportion of variance in pupil academic achievement (ACHI) explained by each variable in the model is shown in *Table 7* and *Figure 3*. These data demonstrated the relative importance of each factor in influencing pupil academic achievement. The proportions of variance of pupil academic achievement presented in *Table 7* and *Figure 3* were obtained through analyses of a series of models of pupil achievement. Variables were sequentially added to the model. Those variables most distant from pupilacademic achievement (ACHI) (variables located to the very left of the model presented in *Figure 3*) were added first. The more proximal variables were added later. This was done to observe how the proportion of variance in pupil academic achievement (ACHI) explained by each model was changed when a new variable was introduced to the model, as the change in the proportion of variance in the dependent variable explained by the model when a new variable is added is the proportion of variance of the dependent variable explained by that variable (Bryman and Cramer, 1997; Pedhazur and Schmelkin, 1991).

Table 7 and Figure 3 about here

It can be seen from *Table 7* and *Figure 3* that together all variables in the model explained 32 per cent of the variance in pupil academic achievement (ACHI). The largest proportion of variance in pupil academic achievement (ACHI) was explained by family socio-economic status (ZPSES) (10 %). Smaller proportions of variance in pupil academic achievement (ACHI) were explained by class repetition (ZPREPEAT) and school location (ZSLOC) (6 and 4 % respectively). The proportions explained by speaking English outside school (ZPENGLIS) and total school resources (ZSRTOT22) were both three per cent. The proportions explained by other variables in the model such as pupil-teacher ratio (ZSPTRATI), teacher reading score (ZRALOCT) and regularity of meals (ZPREGME) were small (i.e. 1-2 %). This meant that adding these variables to the model did not cause much change in the proportion of variation of pupil academic achievement (ACHI) that was explained by the model. Consequently, a third model of the most important variables explaining the variance in pupil academic achievement in Botswana was re-specified and re-examined using AMOS.

Final Re-specified Model of the Important Factors Influencing Pupil Achievement in Botswana

The re-specified model of important factors comprised the following observed variables: school location (ZSLOC), family socio-economic status (ZPSES), total school resources (ZSRTOT22), class repetition (ZPREPEAT), pupil maths score (ZMALOCP) and

pupil reading score (ZMALOCP). There was also one latent variable referred to as pupil academic achievement (ACHI). The solution to the re-specified model of important factors influencing pupil achievement in Botswana is presented in *Figure 4* and *Table 8*.

Both the model fit indices and the model comparison fit indices suggested that the re-specified model of important factors fitted the data well. All of the model comparison fit indices such as NFI, CFI, and IFI were 0.99. The model fit index RMSEA was 0.04, and the GFI and AGFI were 1 and 0.99 respectively, indicating a good fit to the data. The model explained 25 per cent of the variance of the latent variable pupil academic achievement (ACHI).

Figure 4 and Table 8 about here

The interrelationships of the four important factors that influenced pupil academic achievement in Botswana (cchool location (ZSLOC), family socio-economic status (ZPSES), class repetition (ZPREPEAT) and total school resources (ZSRTOT22)) can be seen in *Table 8* and *Figure 4*. That is:

- The variable school location (ZSLOC) had only an indirect effect on pupil academic achievement (ACHI) and this effect was via family socio-economic status (ZPSES) and total school resources (ZSRTOT22). From this it can be inferred that when other things were equal, there was no difference in pupil

achievement between urban and rural areas. However, pupils residing in urban areas were more likely to have higher socio-economic status than those living in rural areas. In addition, pupils from urban areas tended to attend schools with more resources than their counterparts in rural areas. In turn, pupils who came from families with higher socio-economic status and whose school had more resources tended to have higher achievement in both maths and reading. Such findings may explain why pupils in rural areas tended to have lower achievement results than their peers in urban areas.

- The variable family socio-economic status (ZPSES) had both direct and indirect effects on pupil academic achievement (ACHI). The magnitude of the direct effect of family socio-economic status (ZPSES) on pupil academic achievement (ACHI was $\beta_{7.1}$=-0.29 (*t*=-16.4). From this it can be inferred that, when other things were equal, pupils with low family socio-economic status were more likely to have lower achievement results than their peers with high family socio-economic status. The magnitude of the indirect effect via total school resources (ZSRTO22) and class repetition (ZPREPEAT) was less important than the direct effect (the magnitude of the indirect effect was one third of the magnitude of the direct effect). However, this indicated that pupils whose families had higher socio-economic status were more likely to go to schools with more resources and were less likely to repeat class than their counterparts from lower socio-economic status families. That is, low levels of

school resources and class repetition contributed to the observation that pupils from low socio-economic backgrounds tended to be less successful in maths and reading than their counterparts.

The finding that family socio-economic status was the most powerful influence on pupil achievement is in line with the recent findings from the PISA study in which it was reported that "*although PISA shows that poor performance in school does not automatically follow from a disadvantaged socio-economic background, socio-economic background does appear to be a powerful influence on performance"* (OECD, 2004:191).

- Class repetition (ZPREPEAT) was influenced by family socio-economic status (ZPSES) and school location (ZSLOC). In turn, class repetition (ZPREPEAT) had a significant direct effect on pupil academic achievement (ACHI) ($\beta_{7.3}$=-0.27, *t*=-16.6). This suggested that pupils who repeated class were behind their peers in both maths and reading.

- Total school resources (ZSRTO22) was also influenced by family socio-economic status (ZPSES) and school location (ZSLOC). This suggested that pupils from urban areas and from high socio-economic background tended to go to school with more resources. In turn, total school resources (ZSRTO22) directly influenced pupil academic achievement (ACHI) ($\beta_{7.9}$=0.19, *t*=11.2). That is, even when other things were equal, the amount of school resources had an effect on pupil achievement. Pupils from schools with more resources were

more likely to achieve better results than those who attended schools with fewer resources.

Such findings were consistent with the study conducted by Griffin et al. (2003) on monitoring achievement levels in Vietnam, where it was reported…..

> *Compared with pupils from rural or urban areas, pupils from isolated areas achieved considerably lower mean scores in both reading comprehension and mathematics, but these analyses also revealed an influence of access to books that was independent of school location.*
>
> (Griffin, et al., 2003:7)

Policy Implications for Botswana

For any country, it is important to know which factors to concentrate upon when planning initiatives to improve pupil academic achievement. The results of this investigation showed similarities to other international findings regarding the factors that influence pupil achievement, such as PISA (OECD, 2004) and the monitoring of pupil achievement in Vietnam (Griffin et al., 2003). This study has shown that, by applying SEM to the SACMEQ II archived data set for Botswana, the most important factors influencing pupil academic achievement within that country could be determined and used to inform policy planning

(refer to *Figure 4*). However, the simple model of factors influencing pupil achievement (refer to *Figure 2*) provided an extended list of factors to which policy planners within Botswana should also pay attention. In order to provide intervention strategies with both focus and breadth, it was decided to discuss implications based on both the re-specified models referred to in *Figure 2* and *Figure 4.*

As school location (ZSLOC), family socio-economic status (ZPSES), class repetition (ZPREPEAT) and total school resources (ZSRTO22) were important factors associated with pupil academic achievement (ACHI) in Botswana, it might be inferred that policies designed to improve pupil academic achievement would be most effective if applied in association with these factors. However, it is important to distinguish the factors that could be changed and the factors that cannot be changed. It is also crucial to indicate long-term solutions and immediate solutions. In addition, it is important to highlight mediating factors, as a change in those factors reduces the effect of other factors. From the four important factors, both school location (ZSLOC) and family socio-economic status (ZPSES) cannot be changed, at least not in the short-term. In relation to family socio-economic status (ZPSES), while the socio-economic status of people in Botswana can be improved, it is not practical to discuss the influence of the educational sector on the changes in socio-economic status of pupil families. The factor that might be directly malleable is the total school resources (ZSRTO22). This factor was also found to have had a direct effect on pupil academic achievement (ACHI) and acted as a mediating factor between school location (ZSLOC) and family socio-economic status (ZPSES) and pupil academic achievement (ACHI). Improving total school resources

(ZSRTOT22) would not only led to a reduction of the effects of these factors on pupils' achievement, but would also lessen the indirect effect of other factors. This raised the question as to which schools should be targeted. The findings indicated that rural schools and schools with pupils with low family socio-economic status were also the ones with the lowest total school resources. Hence, providing rural schools with more resources would help to reduce the gap in pupil academic achievement of rural and urban areas and of different family socio-economic status levels.

In the United States in the early 1960s, the philosophy that schools should be actively involved in the process of changing society and promoting equality, rather than being controlled by the surroundings and merrily adapting to them, was popular (Cremin, 1961). Thus, compensating education programs were designed to overcome disadvantages associated with poor socio-economic origins (Levin, 1972 cited in Rumberger, 1983). In Australia, similar measures were taken to encourage and provide better opportunities for pupils from disadvantaged backgrounds to receive a good education. The experiences from the US and Australia could provide a model for Botswana in developing pilot development projects to compensate for the disadvantages faced by pupils with financial difficulties, and pupils whose parents have little education. This factor should be regarded as an indicator of disadvantage, based upon which schools should be allocated more resources and funding. The "compensatory programs" may not totally equalize educational opportunity (Levin, 1972 cited in Rumberger, 1983), but the resources would give schools the potential to develop programs and plans to improve academic achievement of pupils. However, in order to provide more

detailed recommendations, further investigation should focus on collecting and introducing the type and quality of resources into the causal model specified in this paper (as opposed to simply the number of resources) to examine the ways in which such factors influence pupil academic achievement.

Recommendation 1: Future research needs to be undertaken to examine the relationship between the *type* and *quality* of resources and pupil academic achievement to assist with identifying appropriate intervention strategies.

As pupils who repeated class at least once were found to have lower achievement levels than their peers, measures should also be taken to help reduce this gap. Although family socio-economic status did not have a strong influence on class repetition, it was found that pupils with low socio-economic status were more likely to repeat class than their higher socio-economic status counterparts. As such, repeating pupils from low socio-economic status should be given more attention.

Recommendation 2: As suggested by Keitheile and Mokubung (2005), it is recommended that further research be undertaken to further investigate the problem of repetition as well as its causes, to enable appropriate intervention strategies to be implemented

The effect of family socio-economic status can also be reduced if other mediating factors such as class repetition, speaking English outside school, regularity of meals and pupil–teacher ratio were improved (see *Table 6* and *Figure 2*). Although the effects of these factors were

not as important as those included in *Figure 4* and *Table 8*, interventions targeting these factors would help to reduce the direct effects of these factors, as well as the effects of family socio-economic status and school location mediating through them.

In relation to regularity of meals, although the meal index for pupils in Botswana was 10.7 (which is considered high) possibly due to the fact that schools in Botswana have implemented a school feeding program (Keitheile and Mokubung, 2005), the results indicated that more effective measures are needed to help pupils suffering from low nutrition.

Recommendation 3: More effective measures should be implemented to provide extra meals for pupils from low family socio-economic status. A study of nutrition among primary school pupils should be undertaken to identify appropriate points of intervention.

As teacher reading score was found to contribute to the variance in pupil achievement in Botswana, programs need to be developed that target teachers' improvement in their knowledge and skills in reading. However, as neither the pupils' reading nor mathematical scores were significantly correlated with the number of in-service programs their teachers had attended (refer to *Table 2*), then professional development programs need to be specifically designed and providers made accountable for raising teachers' skills and knowledge in reading literacy.

Recommendation 4: Targeted professional development programs need to be developed and implemented to increase teachers' knowledge and skills in reading literacy. Providers of such

professional development programs need to be made accountable for the outcomes achieved by participants.

Whilst pupils may be encouraged to speak English at home, this may not always be successful or logistically possible within some family structures. As such, additional tutorials should be designed to help pupils who do not speak English at home to compensate for those who have such opportunities. Whilst coming from a non-English background may not necessarily explain low achievement in numeracy (Rothman and McMillan, 2003), opportunities to practise the assessment language at home have been found to improve educational achievement levels. For example, in the 2004 PISA study, it was reported that…

….pupils who do not speak the language of assessment at home in Belgium, Germany, the Netherlands and Switzerland are at least 2.5 times more likely to be in the bottom quarter of mathematics performance. More generally, being a non-native pupil or speaking a language at home that is different from the language of assessment has a negative impact on mathematics performance, on average across OECD countries, 19 of 9 score points respectively.

(OECD, 2004:169).

Countries like Australia with a continuing flow of migrants are very experienced in this kind of support.

Recommendation 5: Measures such as extra tutorials in English should be provided to help pupils who do not speak English outside school. Teachers' qualifications in this area should also be surveyed to identify the teaching needs in the learning area of English as a second language.

Furthermore, as the analyses showed that pupils were more likely to attain higher achievement where the pupil–teacher ratio was lower, the government may consider ways of reducing the pupil–teacher ratio. Schools in urban areas were more likely than schools in rural areas to be suffering from a high pupil–teacher ratio.

Recommendation 6: Recruitment programs are needed to increase the supply of teachers. Pupil–teacher ratio should be reduced, especially for large schools in urban areas.

Whilst these policy strategies have been presented according to the relative importance to improving pupil achievement as determined by the SEM findings, further analysis on the cost, time and effectiveness of such policies would need to be undertaken by the educational policy makers and planners in Botswana to determine the most feasible and cost effective strategies for improving pupil achievement levels in both the short and long term future.

References

Arbuckle, J. L. (2004). *AMOS* (Version 5.0.1) [Computer software]. Chicago: SmallWaters Corporation.

Bollen, K.A. (1989). Structural equations with latent variables. New York: Wiley.

Bryman, A., and Cramer, D. (1997). Quantitative data analysis with SPSS for Windows : a guide for social scientists. London ; New York: Routledge.

Cremin, L.A. (1961). The transformation of the school: progressivism in American education 1876-1957. New York: Vintage Books.

Falayajo, W., Makoju, G., Okebukola, P., Onugha, D., and Olubodun, J. (1997). *Assessment of learning achievement of primary four pupils in Nigeria*, National report. Federal Ministry of Education Nigeria/UNICEF/UNESCO.

Griffin, P. (2000). A *Global Analysis of achievement in Developping Nations*. Paper prepaed for UNICEF presentation in Dakar, May 2000

Griffin, P., and Tran, K. (2000). *The World Bank Project report on primary schools monitoring*, Ministry of Education and Training, Ha Noi.

Griffin, P., Woods, K., Dulhunty, M., Nguyen, K.C., and Calvitto, L. (2003). Vietnam Pupil Achievement National Assessment of Mathematics and Reading Comprehension Performance of Grade Five Pupils and their Teachers Supplementary Analyses, Melbourne: The University of Melbourne.

Keeves, J.P. (1972). Educational environment and student achievement: a multivariate study of the contributions of the home, the school and the peer group to change in mathematics and science performance during the first year at secondary school. Melbourne: Australian Council for Educational Research.

Keitheile, M., and Mokubung, M., (2005). *The Quality of Primary Education: Some Policy Suggestions Based on a Survey of Standard 6 Pupils*. Botswana: Ministry of Education.

Kulpoo, D. (1998). *The quality of education: some policy suggestions based on a survey of schools.* SACMEQ Policy Research No.1: Mauritius. Paris: IIEP

Milner, Chimombo, Banda and Mchikoma (2001). *The quality of education. some policy suggestions based on a survey of schools.* SACMEQ Policy Research No.7: Malawi Paris: IIEP.

Nassor, S.,& Mohammed, K.A. (1998). *The quality of education: some policy suggestions based on a survey of schools.* SACMEQ Policy Research No.4: Zanzibar. Paris: IIEP.

Nguyen, K. C. (2002). Family and student influences on withdrawal from lower secondary school in rural Vietnam. Unpublished Ph.D, The University of Melbourne, Melbourne.

Nguyen, K. C. (1998). *The relationships between gender, students' achievement and dropping out*. Masters thesis. Victoria University of Technology, Melbourne

Nguyen, K.C., and Fahey, S. (2001). The relationships between gender and students' achievement in rural Vietnam. *Melbourne Studies in Education 42*: 49-68.

OECD (2004). How student performance varies between schools and the role that Socioeconomic Background Plays in this: Learning for Tomorrow's World – First results from PISA 2003. Paris: OECD.

Pedhazur, E.J., and Schmelkin, L.P. (1991). *Measurement, design, and analysis : an integrated approach.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Postlethewaite, N. (1997) *Educational research: some basic concepts and terminology*, International Institute for Educational planning, Paris: UNESCO.

Rosier, M.J. (1978). Early school leavers in Australia : family, school and personal determinants of the decision of 16-year-old Australians to remain at school or to leave: Almqvist and Wiksell International; Australian Council for Educational Research.

Ross, K., Saito, M., Dolata, S., Ikeda, M., and Zuze, L. (2004). Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) Data Archive for the SACMEQ1 and SACMEQ II Projects Version 4.0 (October, 2004) User's Guide. Paris: UNESCO.

Rothman, S., McMillan, J. (2003). Influences on Achievement in Literacy and Numeracy, Longitudinal surveys of Australian youth, Research report Number 36. Melbourne: Australian Council for Education Reasearch.

Rothman, S. (2002). Achievement in Literacy and Numeracy by Australian 14 year-olds, 1975-1998., Longitudinal surveys of Australian youth, Research report Number 29. Melbourne: Australian Council for Education Reasearch.

Rumberger, R.W. (1983). Dropping out of high school: the influence of race, sex and family background. *American Educational Research Journal 20*: 199-220.

Rumberger, R.W. (1995). Dropping out of middle school - a multilevel analysis of students and schools. *American Educational Research Journal 32*: 583-625.

Rumberger, R.W., Ghatak, R., Poulos, G., Ritter, P.L., and Dornbusch, S.M. (1990). Family influences on dropout behavior in one California high- school. *Sociology of Education 63*: 283-99.

Schumacker, R.E., and Lomax, R.G. (1996*). A beginner's guide to Structural Equation Modelling*. Mahwah: Lawrence Erlbaum Associates.

Smith, K.H. (2000). The self-concept and verbal academic achievement of primary and secondary student teachers. Ph.D thesis. The University of Melbourne, Melbourne.

Steiger, J.H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research 25*: 173-80

Tilley, A. J. (1993). An introduction to psychological research and statistics. Pineapple Press.

Tran K. (1994). The situation of repetition and dropping out of students in general schools in Vietnam from 1981 to 1990. In *National Project reports on dropping out and repetition*, ed. K Tran, TT Truong, pp. 1-10. Viet Nam: National Research Institute for Educational Sciences

Murimba S., Moyo G., Pfukani P., Machingaidze T. and Mtembo R. (1994) *The Analysis of Educational Reseach Data for Policy Development: An example from Zimbabwe*, International Institute for Educational Planning,Paris: UNESCO

Williams, T., Long, M., Carpenter, P.H., and Hayden, M. (1993). *Year 12 in the 1980s*. Canberra: AGPS.

Willms, J.D., and Somers, M.A. (2001). Family, classroom, and school effects on children's educational outcomes in Latin America. *School Effectiveness and School Improvement 12*: 409-45.

World Bank. (2004). Vietnam, reading and mathematics assessment study, Volume 2. Human Development Sector Unit East Asia and Pacific Region.

Table 1

Pupil, home, teacher and school background variables extracted for further analyses using correlation analyses.

Variable Type	Variable Name	Variable Code	Descriptions and Values

Pupil Achievement Related	Maths score	ZMALOCP	Pupil Math-all 500 score [mean=500 and SD=100]
	Reading score	ZRALOCP	Pupil Reading-all 500 score [mean=500 and SD=100]
Pupil Background Characteristics	Pupil gender	PSEX	Pupil gender (1 = male; 2 = female)
	Pupil age	ZPAGEMON	Pupil's age in months rounded to the nearest
	Absence	PABSENT	Pupil number of days absent
	Class repetition	ZPREPEAT	Pupil repeated class (0= never, 1= repeated at least once)
Home Characteristics	Family SES	ZPSES	Pupil's SES [parents' education; possessions at home; light; wall; roof, floor]. Maximum was 15.
	Regularity of meals	ZPREGME	Pupils were asked to indicate whether for each of the breakfast/morning meals (PMEAL1), lunch (PMEAL2) and evening meals (PMEAL3) how often they normally eat (1=not at all; 2=1or2/week; 3=3or4/week; 4=every day). This is the derived variable of these three variables ZPREGME = PMEAL1 + PMEAL2 + PMEAL3 (Range 3-12)
	Home interest composite	ZPHMINT	This is a composite variable that measures people from home showing interest in pupil learning. The scoring for this variable is the sum of the scoring of the five variables ZPHMINT = ZPHMWKDO + ZPHMWKHE + ZPASK + ZPQUEST + PLOOKWK. Ranged from 5 to 15.

	Speaking English outside school	ZPENGLIS	Speaking English outside school (1=never, 2=sometimes, 3 = often, 4=all the time)
Teacher Characteristics	Reading teacher qualification	ZXQPROF	Reading teacher training level
	Maths teacher qualification	ZYQPROF	Math teacher training level
	Teacher maths score	ZMALOCT	Teacher Math-all 500 score [mean=500 and SD=100]
	Teacher reading score	ZRALOCT	Teacher Reading-all 500 score [mean=500 and SD=100]
	Reading teacher years of teaching	XEXPER	Reading teacher years of teaching
	Maths teacher years of teaching	YEXPER	Maths teacher years of teaching
	Reading teacher PD courses	XINSERVC	Reading teacher number of in-service courses attended
	Maths teacher PD courses	YINSERVC	Maths teacher number of in-service courses attended
School Location	Pupil-teacher ratio	ZSPTRATI	Pupil-teacher ratio

and Class and School Resources	Reading total class resources	ZXCLRES8	Reading teacher total class resources [max=8]
	Maths total class resources	ZYCLRES8	Math teacher total class resources [max=8]
	Total school resources	ZSRTOT22	Total school resources [max=22]
	School location	ZSLOC	School location (1=isolated/ rural, 2=small town or large town/city).

Table 2

Correlations between selected background variables and pupil achievement in Botswana at the end of primary school.

Factor	Variable	Pupil Maths Score	Pupil Reading Score
Pupil	Pupil gender	0.06	0.15
	Pupil age	-0.23	-0.28
	Absence	-0.08	-0.09
	Class repetition	-0.27	-0.27
Family	Socioeconomic status	0.30	0.39
	Regularity of meals	0.17	0.21
	Home interest composite	0.08	0.13
	Speaking English outside school	0.22	0.25
Teacher	Reading teacher qualification	0.04	0.06
	Maths teacher qualification	0.04	0.08
	Teacher maths score	0.13	0.14
	Teacher reading score	0.18	0.18
	Reading teacher years of teaching	0.11	0.12
	Maths teacher years of teaching	0.08	0.10
	Reading teacher PD courses	0.03*	0.04*

	Maths teacher PD courses	0.01*	0.02*
School Location and Resources	Pupil-teacher ratio	-0.17	-0.15
	Reading total class resources	0.04*	0.02*
	Maths total class resources	0.04*	0.02*
	Total school resources	0.29	0.32
	School location	0.15	0.21

Note: * not significant at the p<0.01 level.

Table 3

Pupil, home, teacher and school background variables extracted for further analyses using SEM

Factor	Variable name	Variable code	Variable number
Pupil Academic Achievement	Maths score	ZMALOCP	V_6
(ACHI, V_7)	Reading score	ZRALOCP	V_5
Pupil Background	Pupil age	ZPAGEMON	V_4
	Class repetition	ZPREPEAT	V_3
Home Background	Socioeconomic status	ZPSES	V_1
	Regularity of meals	ZPREGME	V_8
	Speaking English outside school	ZPENGLIS	V_2
Teacher Characteristics	Teacher reading score	ZRALOCT	V_{11}
School/Class Characteristics	Pupil-teacher ratio	ZSPTRATI	V_{10}
	Total school resources	ZSRTOT22	V_9
	School location	ZSLOC	V_0

Figure 1. Hypothesised saturated model of factors influencing pupil academic achievement in Botswana

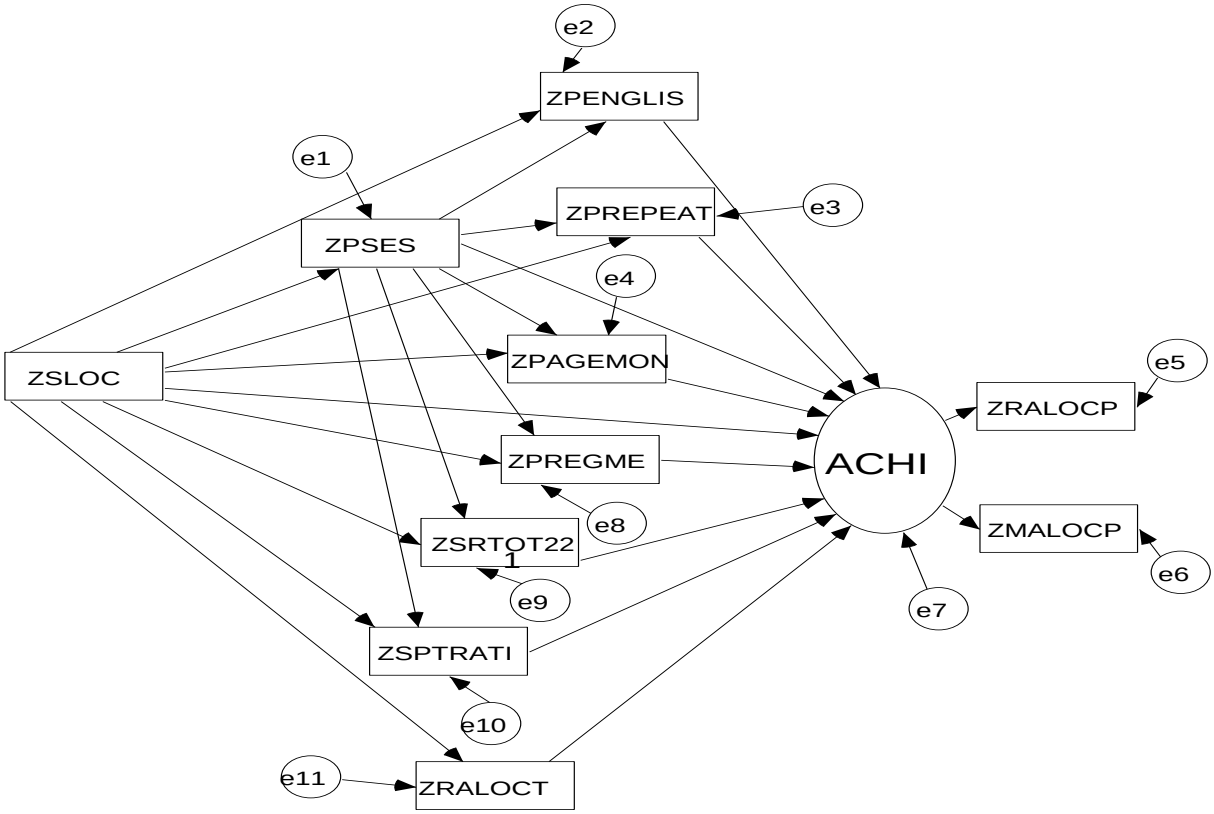

Table 4

Extracted variance and measurement error of observed variables of the latent construct achievement (ACHI)

Measurement Observed Variables	Factor Loading	Extracted Variance	Residual Variance
Pupil reading score (ZRALOCP)	0.94	88%	12%
Pupil maths score (ZMALOCP)	0.82	67%	33%

Table 5

Direct, indirect and total effects of factors on pupil achievement – Botswana saturated model

Variable	Direct Effect	Indirect Effect	Total Effect
School Location on:			
Family socioeconomic status	0.37	--------	0.37
Pupil age	-0.07	-0.1	-0.17
Class repetition	-0.01*	-0.04	-0.05
Teacher reading score	0.06	--------	0.06
Regularity of meals	-.0.05*	0.07	0.02
Speaking English outside school	0.07	0.08	0.15
Total school resources	0.19	0.12	0.31
Pupil-teacher ratio	0.22	-0.03	0.19
Pupil achievement	0.06*	0.15	0.21
Family Socioeconomic Status on:			
Pupil age	-0.25	--------	-0.25
Class repetition	-0.11	--------	-0.11
Regularity of meals	0.2	--------	0.2
Speaking English outside school	0.21	--------	0.21
Total school resources	0.32	--------	0.32
Pupil-teacher ratio	-0.09	--------	-0.09
Pupil achievement	0.2	0.16	0.36
Speaking English Outside School on:			
Pupil achievement	0.15	--------	0.15
Class Repetition on:			
Pupil achievement	-0.23	-------- -	-0.23

Pupil Age on:			
Pupil achievement	-0.08	--------	-0.08
Regularity of Meals on:			
Pupil achievement	0.14	--------	0.14
School Total Resources on:			
Pupil achievement	0.14	--------	0.14
Pupil - Teacher Ratio on:			
Pupil achievement	-0.13	--------	-0.13
Teacher Reading Score on:			
Pupil achievement	0.12	--------	0.12

Note: * Effects not significant beyond p<0.05 level; dashed lines (-----) represent zero effects.

Table 6

Direct, indirect and total effects of factors on pupil academic achievement – The re-specified simple model for Botswana

Variable	Direct Effect	Indirect Effect	Total Effect
School Location on:			
Family socioeconomic status	0.37	------	0.37
Class repetition	------	-0.04	-0.04
Regularity of meals	------	0.07	0.07
Speaking English outside school	------	0.09	0.09
Total school resources	0.19	0.12	0.31
Pupil-teacher ratio	0.22	-0.03	0.19
Pupil achievement	------	0.14	0.14
Family Socioeconomic Status on:			
Class repetition	-0.12	------	-0.12
Regularity of meals	0.18	------	0.18
Speaking English outside school	0.24	------	0.24
Total school resources	0.32	------	0.32
Pupil-teacher ratio	-0.09	------	-0.09
Pupil achievement	0.23	0.15	0.38
Speaking English Outside School on:			
Pupil achievement	0.15	------	0.15
Class Repetition on:			
Pupil achievement	-0.26	------	-0.26
Regularity of Meals on:			
Pupil achievement	0.13	------	0.13
School Resources on:			

Pupil achievement	0.16	------	0.16
Pupil-teacher ratio on:			
Pupil achievement	-0.12	------	-0.12
Teacher Reading Score on:			
Pupil achievement	0.12	------	0.12

Note: * Effects not significant beyond $p<0.05$ level; dashed lines (-----) represent zero effects.

Figure 2. Solution to the re-specified simple model of factors influencing Pupil Academic Achievement in Botswana

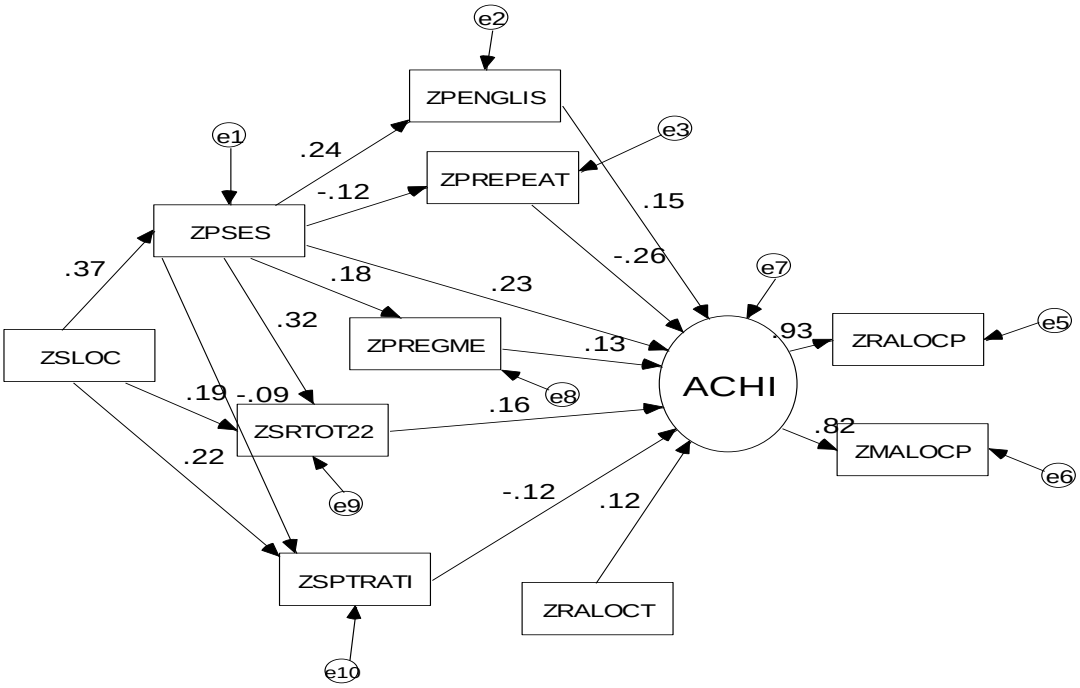

Table 7

Variance in achievement (ACHI\) explained by each variable in the Botswana re-specified simple model

Variables	Proportion of Variance in Achievement (ACHI) (%)
School location	4
Family socioeconomic status	10
Regularity of meals	2
Speaking English outside school	3
School total resources	3
Teacher reading score	1
Pupil-teacher ratio	1
Class repetition	6
Model presented in Figure 2	30

Figure 3. Proportion of variance in achievement (ACHI) explained by each variable in the Botswana re-specified simple model.

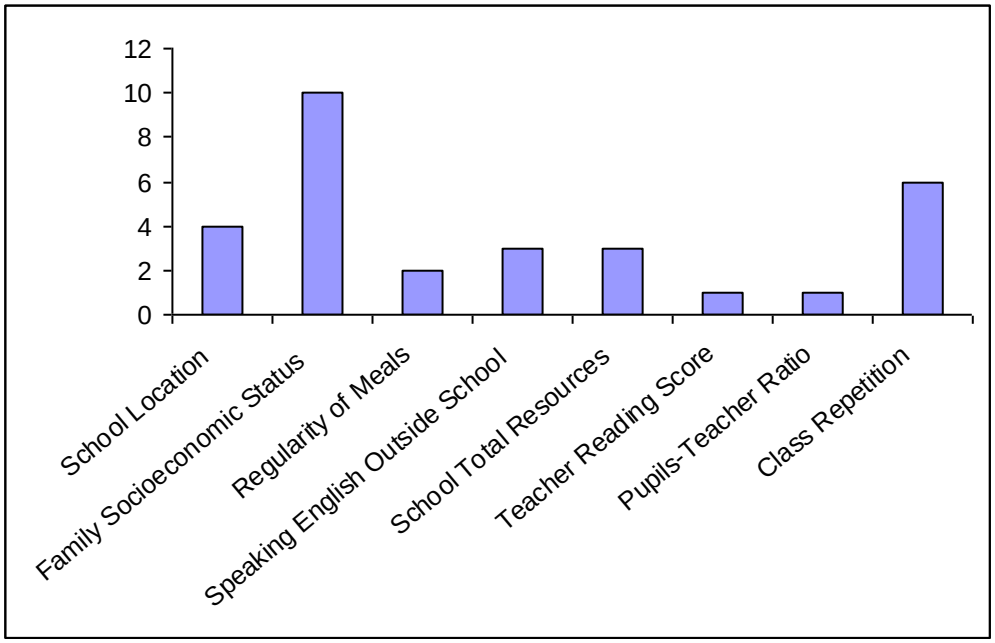

Table 8

Direct, indirect and total effects of factors on pupil achievement -Solution to the final re-specified model of important factors influencing pupil achievement in Botswana.

Variable	Direct Effect	Indirect Effect	Total Effect
School Location on:			
Family socioeconomic status	0.37	------	0.37
Total school resources	0.19	0.12	0.31
Class repetition	------	-0.04	-0.04
Pupil achievement	------	0.17	0.17
Family Socioeconomic Status on:			
Total school resources	0.32	------	0.32
Class repetition	-0.12	------	-0.12
Pupil achievement	0.29	0.09	0.38
Class Repetition on:			
Pupil achievement	-0.27	-------	-0.27
Total School Resources on:			
Pupil achievement	0.19	-------	0.19

Note: Dashed lines (-----) represent zero effects

Figure 4. Solution to the final re-specified model of important factors influencing pupil achievement in Botswana.