

Stability and Consistency of School Effects and the Implications for School Improvement Interventions: The Case of Botswana

I. L. (Irene) Mohiemang¹ and S. G. Pretorius²

¹*Maun Homeopathy Project, Maun, Botswana
E-mail: imohiemang@yahoo.co.uk*

²*Department of teacher Education, University of South Africa, P O Box 392,
Pretoria 0003, South Africa
Telephone: 27124600969, Fax: 27866343903, E-mail: pretosg@unisa.ac.za*

KEYWORDS School Effectiveness. School Effects. School Improvement. Value Addition. Multilevel Modelling

ABSTRACT School effectiveness is not a stable and consistent school characteristic. The aim of this study was to determine the consistency of departmental effects within secondary schools of Botswana and the stability of school and departmental effects across the years. A value added methodology, controlling for prior attainment and gender was employed. *MLwiN* 2.10 Beta (4), based on multilevel modelling, analyzed the Botswana General Certificate of Secondary Education (BGCSE) examination results from a sample of 5662 candidates for 2005-2007. There was evidence of departmental variation within the same school. The important implications for departmental and school improvement as well as for reporting school performance are discussed.

INTRODUCTION

School effectiveness is defined empirically as the degree to which schools achieve their goals (Scheerens 2000). The goal of schools is to educate all learners in all areas of the curriculum. School Effectiveness Research (SER) regards the role of the school as adding value to

learners' prior attainment across all subject areas. One area that researchers in SER are faced with is the extent to which schools achieve their goal of educating learners across all the subjects in a school. Schools are dynamic organizations that keep on changing leaders, teachers and cohorts. The challenge that comes with these changes is whether schools' performances remain the same or not.

The unidimensionality of the school concept is considered to be one of the fundamental issues in SER (Scheerens and Bosker 1997). Unidimensionality refers to whether the results of schools and classes are the same for different subjects, school years and groups of students (Kyriakides and Creemers 2008). Consistency and stability are operationally defined as the correlation between two rank orderings of schools (Scheerens and Bosker 1997). Consistency refers to the correlation between rank orderings of schools based on different criterion variables while stability has to do with the extent to which the rank order of schools/departments on outputs remains the same

regardless of the time-point at which the effect is measured (Kyriakides and Creemers 2008). For example, schools may be rank ordered on the basis of their performance this year and then compare this with the rank order for the preceding year (stability) or one might rank them on the basis of their output in English Language and correlate this with their output in mathematics (a consistency measure). The term school effects as it is commonly used by researchers into educational effectiveness, relates to differences between schools regarding student achievement scores that have been adjusted for cognitive aptitude (Luyten 2003).

Schools change administrators, teachers and students. To compare cohorts of students and to compare the performance of schools based on the proportion of those obtaining five or more A*-C grades from one year to another is not enough and is inappropriate because performances at intakes differ from one year to another. School performance indicators based on the credit passes (therein after, raw results) have been criticized heavily by school effectiveness researchers and have been superseded by value added methods. SER has developed from correlates of effective schools through school improvement to value added methodology using sophisticated statistical techniques (Peng et al. 2006). These latter developments have led to more accurate methods of evaluating school performance. One of the key methodological developments is multilevel modelling and has been used to

calculate educational effectiveness within (departments) and between schools in terms of the value added and to scrutinize in more detail SER fundamental issues such as the true size and stability over time of school effects (Peng et al. 2006). Sammons (2006) notes that longitudinal datasets are vital for the detail exploration of teacher, class, or school effects.

The consistency and stability of school effects are one of the most fundamental and complex issues that face researchers within the field (Schreens and Bosker 1997; Luyten 1998). Scholars in educational effectiveness note that studies on the stability of school effects have important implications for theory development within the field (Kyriakides and Creemers 2008) and only international studies can tap the full range of variation in the school and classroom quality, and therefore in potential school and classroom effects (Kyriakides 2006). The research will make a contribution to the development of theory on the important issues of SER.

Different stakeholders in schools need to know whether school effects on students are consistent and stable. Parents and children need the information for school choice. From a policy standpoint, if there is a policy to reward effective schools, when the consistency is not known, ineffective subject departments in the effective schools will benefit. Similarly, effective departments in ineffective schools can be demoralized since they will never be rewarded.

Ma (2000) argues that consistency and stability in learning outcomes are becoming a type of equity issue. If students are to achieve their full academic potential, then it becomes a disaster if they progress well in some subjects and do not progress in others. This imbalance is considered to be lack of equity. Schools change administrators, cohorts and teachers. Effective schools or teachers are expected to change and adopt as their context change. This is consistent with the contingency theory and the main models of educational effectiveness (Scheerens and Bosker 1989).

The current research is particularly needed given that despite the fact that value added approaches that seek to identify and measure the individual school's contribution to student progress have become increasingly recognized as providing the most appropriate methodology to explore school effects (Sammons 2006). No study has been carried out to investigate the stability and consistency of senior secondary school performance in the BGCSE using a value added

methodology. In addition, the performance of senior secondary schools is based on the proportion of students who have obtained five or more A*-C grades at BGCSE. Every year the performance of cohorts is compared subject by subject and in overall performance. Judging school performance based on this is problematic because of the changes in cohort intake abilities from year to year.

Although the overall effectiveness of senior secondary schools for the years 2005-2007 has been identified in a separate study (Mohiemang 2009), there might be some considerable variation between subject departments in schools and across cohorts. Therefore, the main purpose of this study is to examine whether school effects in the sampled schools are consistent and stable.

The following research questions were addressed:

- ♦ Do schools perform consistently across subject departments?
- ♦ Are school effects stable over time (across cohorts)?

Before presenting the results of this study, previous research on consistency and stability of school effects is presented to contextualize the study.

Theoretical and Empirical Background

Classical bureaucracies in organizations under laid the thinking on early studies of SER where effectiveness was considered to be a consistent and stable school characteristic (Luyten 1994). The early studies on educational effectiveness assumed schools as classical bureaucracies where effectiveness was considered to be unidimensional (Luyten 1994). When a school was rated as effective, the assumption was that it was effective in all subjects and across the years. This unidimensionality of effectiveness was challenged based on the characterization of schools as "loosely coupled systems" and professional bureaucracies in which teachers are autonomous in their classrooms. From these theoretical considerations, effectiveness cannot be assumed to be a stable school characteristic (Luyten 1998). Schools that are effective in one subject are not always effective in other subjects and stability across cohorts is sometimes problematic (Luyten 1994; Scheerens and Bosker 1997).

Teddlie et al. (2000) note that the results on consistency and stability varied depending on the country and methodology employed. Studies

investigating the stability of school effects revealed that school effects are stable to a certain degree (Luyten 1994; Gray et al. 1995). Earlier studies on consistency (Cuttance 1987; Smith and Tomlinson 1989; Wills and Raundenbush 1989 in Teddlie and Reynolds 2000) concluded that schools that were successful in one subject were also successful in another and across all subjects. However, later studies found that there was less consistency between subjects in secondary schools. There were some effective and ineffective departments coexisting in the same school (Sammons and Nuttall 1993; Sammons et al. 1997; Luyten 1998, 2003) and hence, the unidimensionality of the school effect concept in secondary school is questionable (Kyriakides 2004). On the basis of these results, Fitz-Gibson concluded that because value added results from different subjects showed moderate correlations, performance in secondary schools should be by subject (Teddlie et al. 2000; Highfield 2010).

The importance of research into the consistency of school effectiveness has been acknowledged by leading scholars in the field. They are vital for further theory development (Scheerens and Bosker 1997; Kyriakides 2004; Kyriakides and Creemers 2008). The results from this study will make contributions to theory development in educational effectiveness. Thomas et al. (1997) argue that the review of previous research on consistency and stability highlights the need for further longitudinal research to address these issues on both overall examination results and performance in specific subjects.

METHOD

The target population for this study was the Form Five (Grade 12) students of all twenty-seven government and government aided secondary schools of Botswana who wrote the BGCSE examinations in 2005, 2006 and 2007. The national datasets for the JCE (Junior Certificate Education) (Form 3) and BGCSE examination results, which were available at the time of conducting the research, were obtained from the Department of Secondary Education and BEC (Botswana Examination Council) respectively. To ensure representation, proportionate simple random sampling, based on probability sampling selected students for the study.

Statistical Analysis

The data collected had a hierarchical structure where students are in departments within years within schools and schools located in inspectorial regions. From a statistical perspective, such data requires the use of the statistical techniques of multilevel modelling (Goldstein 1999) which should incorporate the hierarchical structure of data. The data was analysed via the statistical software package MLwiN 2.10 Beta (4), which is based on multilevel modelling (Rasbash et al. 2008). Four measures of performance were employed, total score and three core subjects in English Language, Setswana Language (mother tongue) and mathematics.

A more general multilevel model called a random coefficient regression model where both the intercept and the slope were random variables, as in Model (1) was used in the analysis where pupils are (indexed by i) at level one and schools (indexed by j) are at level two (Rasbash et al. 2008).

Model 1: A random coefficient model of school effectiveness

$$TBGCSE_{ij} = \beta_{0j} + \beta_{1j} TJC_{ij} + e_{ij}$$

$$\beta_{0j} = \beta_0 + u_{0j}$$

$$\beta_{1j} = \beta_1 + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \sigma_{u0}^2 & \\ & \sigma_{u1}^2 \end{bmatrix}$$

$$e_{ij} = N(0, \sigma_e^2)$$

where

$TBGCSE_{ij}$ is the final BGCSE exam score of student i in school j

TJC_{ij} is the intake /prior attainment score of student i in school j

β_{0j} represents the intercept for school j and can be written as $\beta_{0j} = \beta_0 + u_{0j}$

where β_0 is the average intercept for all the schools in the sample, u_{0j} is the unique effect of school j on the mean of student BGCSE achievement

β_{1j} represents the slope for school j can be written as

$\beta_{1j} = \beta_1 + u_{1j}$ where β_1 is the averaged effect of prior achievement on BGCSE achievement,

u_{1j} is the unique effect of school j on the slope of the regression line

e_{ij} is the individual effect of student i in school j

The terms u_{0j} and u_{1j} are random departures from β_0 and β_1 or residuals at the school level follow a multi variate or a bivariate normal distribution with mean vector 0 and covariance matrix Ω_u . The elements of Ω_u are

$Var(u_{0j}) = \sigma^2_{u0}$ (the variation in the intercepts across the schools' summary line)

$Var(u_{1j}) = \sigma^2_{10}$ (the variation in the slopes across the schools' summary line)

$Cov(u_{0j}, u_{1j}) = \sigma_{01}$ (the covariance between the schools' intercepts and slopes)

The students' scores depart from their school's summary line by an amount e_{ij} , which is assumed to be normally distributed with mean 0 and a variance σ^2_e .

For the value added by schools, the response variables used for each student were the BGCSE average and the explanatory variables were the JCE average score.

Value Added by Subject Department

The value added by subject departments was computed using Model 2. The average for the subject at BGCSE is the dependent variable and the average for the corresponding subject at JCE is the explanatory variable.

Model 2: The value added by subjects

$$Ebgcse = \beta_{0j} + \beta_{1j} engjc_{ij} + e_{ij}$$

$$\beta_{0j} + \beta_{0j} + u_{0j}$$

$$u_{0j} \sim N(0, \sigma^2_{u0})$$

$$e_{ij} \sim N(0, \sigma^2_e)$$

Where **Ebgcse_{ij}** is the average for all students *i* in school *j* for any of the subjects and **engjc_{ij}** represents the prior intake score for all the students in school *j* for English Language or any of the subjects. All the other terms were defined in Model 1.

RESULTS

Table 1 represents descriptive statistics on the sample. It shows the proportion of pupils within each group for each background variable. The sample consists of 5662 students of which 52% were female while 48% were male. In terms of intake

ability, 33.4%, 33.9% and 32.1% students were of high, middle and low ability respectively

Overall Consistency across the Departments for 2005-2007

As an example, Model 2 for 2005 Maths becomes:

$$mathsbgcse_i = -1.135(1.905) + 0.880(0.404) mathsjc_i + e_i$$

$$e_i \sim N(0, \sigma^2_e) \sigma^2_e = 0.059(0.016)$$

$$-2 * \loglikelihood = 0.269(27 \text{ of } 27 \text{ cases in use})$$

The average score of all the students at BGCSE was regressed on to the average prior attainments of all students, therefore only one level, at the school is shown. The average for the three years is calculated to give the overall value added by the departments.

To determine the overall consistency by schools over the three years, the average value added by each department over the years was calculated to yield a single score of departmental effectiveness in Table 2.

Six (22%) schools were consistent in effectiveness across all the subjects. Out of the nine schools that were effective in overall, only five had consistent results while four had ineffective and effective departments coexisting. Out of the eighteen schools that were ineffective overall, one was effective in all the subjects while thirteen had effective and ineffective departments and four (15%) were consistently ineffective for all the three subjects. However, the majority of schools 17(63%) were not consistent across the three subject departments. Table 3 shows the internal variation by departments.

Fifteen (56%) schools were effective for English Language while twelve (44%) were not and thirteen (48%) were effective for Setswana while twelve (44%) were effective for Maths and fifteen (56%) were ineffective for Maths.

The correlation for the value added by the subjects is one other way to determine the consistency between departments as is shown in Table 4.

Table 1: Descriptive statistics of the sample

Year	Female (%)	Male (%)	High ability (%)	Mid ability (%)	Low ability (%)	Total (%)	Population
2005	959 (51.8)	891 (48.2)	614 (33.2)	658 (35.6)	578 (31.2)	1850	17 855
2006	995 (52.1)	916 (47.9)	624 (32.7)	622 (32.5)	665 (34.8)	1911	18 101
2007	990 (52.1)	911 (47.1)	689 (36.2)	637 (33.5)	575 (30.2)	1901	22 067
Total (%)	2944 (52%)	2718 (48%)	1927 (34.0)	1917 (33.9)	1818 (32.1)	5662 (10)	58 023

Table 2: Overall consistencies across departments for 2005-2007

School	English	Setswana	Maths
1	0.20	0.04	0.05
2	0.12	-0.01	-0.11
3	-0.34	-0.27	-0.41
4	0.14	-0.03	-0.04
5	0.06	0.05	-0.08
6	-0.27	0.06	-0.02
7	0.04	-0.20	-0.14
8	0.09	0.10	0.19
9	0.02	-0.09	-0.25
10	-0.03	-0.25	0.18
11	0.19	0.17	0.29
12	0.12	-0.03	0.20
13	-0.11	0.05	0.08
14	-0.15	-0.22	0.07
15	-0.14	-0.10	-0.17
16	-0.32	0.10	-0.20
17	-0.02	0.03	0.17
18	-0.06	-0.04	-0.30
19	0.11	0.09	-0.06
20	0.27	0.32	0.08
21	-0.001	0.19	-0.16
22	-0.12	-0.005	-0.13
23	0.15	0.27	0.27
24	-0.09	-0.14	0.10
25	0.06	0.40	0.43
26	0.06	-0.25	-0.18
27	0.13	-0.22	0.32

Table 3: Variations within schools

School	Effective for	Ineffective for
2	English	Maths, Setswana
4	English	Maths, Setswana
5	English, Setswana	Maths
6	Setswana	English, Maths
7	English	Maths, Setswana
9	English	Maths, Setswana
10	Maths	English, Setswana
12	English, Maths	Setswana
13	Setswana, Maths	English
14	Maths	English, Setswana
16	Setswana	English, Maths
17	Setswana	English, Maths
19	English, Setswana	Maths
21	Setswana	English, Maths
24	Maths	English, Setswana
26	English	Mathsand Setswana
27	English, Maths	Setswana

Table 4: Correlation coefficients between subjects' value added in 2005-2007

	English	Maths	Setswana
English		0.49	0.36
Maths			0.40

The correlations between the subjects were positive but low and weak. This means that in most

schools the departments were not consistent in their outcomes.

Stability of Overall Value Added across the Years

The value added average indicates that nine (33%) schools were more effective than others. Pupils in these schools progressed further than expected from their initial intake grading. However, the majority of schools eighteen (67%) were ineffective. Students in these schools made less progress than similar students (ability and gender) in other schools.

Table 5: Value added by schools and the average for 2005-2007

School value	Value added 2005	Value added 2006	Value added 2007	Average value added	Inspectional area/region
1	1.20	-0.62	-2.41	-0.61	D
2	0.26	-1.69	0.90	-0.18	C
3	-0.01	-2.58	-3.28	-1.958	E
4	0.54	-0.55	-0.76	-0.26	C
5	0.03	-1.58	-1.12	-0.89	C
6	0.84	-0.22	0.89	0.50	C
7	-1.34	-2.34	-2.55	-2.08	A
8	0.49	-1.09	0.79	0.07	C
9	-0.94	-4.42	0.91	-1.48	A
10	0.88	2.11	-2.37	0.21	A
11	0.92	2.3	3.75	2.33	D
12	2.82	4.17	2.11	3.03	D
13	-2.57	0.13	0.65	-0.60	B
14	0.36	-1.80	-1.65	-1.03	A
15	0.22	-1.72	-1.42	-0.97	E
16	-2.43	-1.63	1.96	-0.70	C
17	-3.21	-2.8	1.55	-1.47	A
18	-2.29	-0.27	-1.75	-1.44	C
19	-2.09	-1.81	-1.32	-1.74	B
20	3.82	5.85	5.47	5.05	C
21	-0.57	1.11	-3.99	-1.16	B
22	-1.27	-2.20	-1.86	-1.78	A
23	4.00	7	5.39	5.46	D
24	-1.90	1.09	-1.73	-0.85	A
25	3.52	5.05	1.75	3.44	C
26	-2.42	-3.40	-4.99	-3.60	A
27	1.178	1.87	5.07	2.71	D

Table 5 indicates that six schools (22%) were stable in adding positive value though the amount added was not consistent. Schools 27 and 11 showed an increase in the value added from year to year while schools 25 and 12 showed a decline. Six (22%) schools were stable in performing below expectation and schools, 3, 7 and 26 were also increasingly becoming ineffective. Fifteen (56%) schools were unstable (positive and negative value added). The correlation for 2005 and 2006 is 0.77;

2005 and 2007 is 0.53 and 2006 and 2007 is 0.61. The correlations are positive but not perfect indicating instability across some schools. The results also indicate that some schools were declining in their effectiveness and this included schools 14, 15, 5, 4 and 1 while school 13 was increasingly becoming effective.

The overall value added measure by schools reported in Table 5 revealed that there were nine effective schools. On the other hand, the departmental effectiveness results (Tables 2 and 3) show that only six schools were effective across all the departments. School 1 which was not effective in overall was effective in the three subjects. Schools 6, 10, 11, 12, and 27, which were effective schools, were differentially effective across the departments. Other schools which were not effective overall, however, were effective in one or two subjects, for example, school 13 for Setswana and English and school 19 for English and Setswana.

Departmental Stability across the Years

Twelve (44%) schools were stable in their departmental effectiveness (Table 6). Of the twelve schools, eight added negative value and four schools were stable in positive value added. Fifteen (56%) schools were unstable in their departmental effectiveness.

Fifteen (56%) schools were stable in their effects for the subjects (Table 7). Nine of the schools had negative effects while six had positive stable effects. Twelve (44%) schools were unstable across the cohorts for Setswana language.

Seventeen (63%) schools were stable from year to year (Table 8). Out of these, ten schools showed negative stability while seven schools showed stable positive value added. Ten (37%) schools show instability across the years.

The correlations within the subjects across the years were 0.35, 0.49 and 0.53 for English, Maths and Setswana respectively. The correlations were from weak to moderate and not perfect. This indicates that subjects were not stable across the cohorts.

DISCUSSION

The findings on consistency are in line with earlier empirical studies on the consistency across different subject areas (Sammons et al.1993; Sammons et al. 1997; Luyten 1998; Ma 2001; Doolaard 2002; Luyten 2003; Jakubowski 2008). Thomas and Mortimore (1996) found strong evi-

Table 6: English department

<i>School</i>	<i>English 2005</i>	<i>English 2006</i>	<i>English 2007</i>
1	0.2	0.31	0.02
2	0.29	-0.10	0.17
3	-0.54	-0.22	-0.26
4	0.02	0.24	0.17
5	-0.05	0.17	0.07
6	-0.29	-0.39	-0.11
7	0.13	-0.15	0.16
8	0.09	0.18	-0.01
9	0.13	-0.05	-0.03
10	-0.05	0.03	-0.07
11	0.53	-0.02	0.05
12	0.13	0.27	-0.04
13	-0.19	-0.11	-0.04
14	-0.02	-0.28	-0.16
15	-0.17	-0.07	-0.17
16	-0.39	-0.34	-0.23
17	-0.08	-0.13	0.16
18	-0.19	-0.03	0.05
19	0.17	0.07	0.07
20	0.41	0.34	0.07
21	-0.11	-0.01	0.10
22	-0.24	-0.22	0.11
23	-0.37	0.30	0.51
24	-0.13	-0.15	0.02
25	0.15	0.11	-0.08
26	0.27	0.12	-0.15
27	0.22	0.13	0.04

Table 7: Setswana department

<i>School</i>	<i>Setswana 2005</i>	<i>Setswana 2006</i>	<i>Setswana 2007</i>
1	0.03	0.24	-0.14
2	0.17	-0.04	-0.16
3	-0.21	-0.51	-0.10
4	-0.003	-0.08	-0.01
5	0.11	0.05	-0.01
6	0.08	0.18	-0.09
7	-0.34	-0.22	-0.04
8	0.37	-0.23	0.17
9	0.08	-0.19	-0.15
10	-0.28	-0.32	-0.16
11	-0.12	0.08	0.54
12	-0.19	0.11	-0.02
13	-0.02	0.08	0.08
14	-0.29	-0.12	-0.26
15	-0.03	-0.19	-0.08
16	0.02	0.1	0.17
17	-0.12	-0.01	0.22
18	0.08	0.02	-0.22
19	0.13	0.07	0.09
20	0.47	0.11	0.39
21	0.07	0.28	0.22
22	0.04	0.10	-0.16
23	0.08	0.42	0.31
24	-0.13	-0.07	-0.21
25	0.4	0.49	0.30
26	-0.2	-0.06	-0.49
27	-0.21	-0.28	-0.18

Table 8: Maths department

<i>School</i>	<i>Maths 2005</i>	<i>Maths 2006</i>	<i>Maths 2007</i>
1	-0.05	0.38	-0.17
2	0.35	-0.36	-0.32
3	-0.48	-0.54	-0.22
4	-0.19	-0.03	0.10
5	-0.22	-0.008	-0.007
6	0.06	-0.02	-0.09
7	-0.29	-0.11	-0.04
8	0.15	0.16	0.27
9	-0.05	-0.59	-0.11
10	0.15	0.26	0.13
11	0.38	0.13	0.36
12	0.33	0.31	-0.04
13	0.14	0.01	0.08
14	0.14	-0.09	0.17
15	-0.09	-0.17	-0.26
16	-0.22	-0.27	-0.11
17	-0.25	0.18	0.13
18	-0.25	-0.41	-0.25
19	-0.23	0.02	0.02
20	0.18	0.20	-0.15
21	-0.26	-0.08	-0.14
22	-0.07	-0.13	-0.19
23	0.16	0.36	0.29
24	-0.16	0.20	0.27
25	0.42	0.40	0.46
26	-0.03	-0.13	-0.37
27	0.40	0.31	0.26

dence that schools were not consistently effective in the two areas they analysed, Mathematics and English, with a correlation of 0.46.

Sammons et al. (1997) and Luyten (1994) believe that the concept of school effectiveness needs to be qualified at the school level to the term school and departmental effectiveness. The use of only a single overall measure of value added may hide these differences that have important messages for departmental self evaluation. Trower and Vincent (in Mayston 2007: 40) came to the same conclusion from their empirical study of pupil value added by General Certificate of Secondary Education results when matched to KS3 prior attainment that "a value added indicator based on an outcome measure which combined all subjects would hide the substantial differences in success between different departments within the school."

The results of the study have implications for departmental and school improvement. Research into departmental effectiveness and improvement has emphasized that the department is the key focus for change within the school (Brown et al. 2000). Hopkins (2001) argues that schools that add value to the learning, progress and attainment of their pupils are consistent in their teaching practice which Blanchard (2002) calls instructional efficacy.

This operates on the belief that difficult students are teachable through extra effort and appropriate techniques, and that they can enlist family support and overcome negating community influences through effective teaching. Such teachers devote more time to academic activities, provide students who encounter difficulties with the guidance they need to succeed and praise their academic achievement.

The subject leader who has a responsibility for a manageable group of people can enable successful change within a group and thus contribute to whole school improvement through her or his direct influence upon the quality of teaching and learning within a department. The leadership approach adopted by the subject leader affects the department's performance. One such approach is distributing leadership within the department (Harris 2001). In overall, an important dimension of the leader is to shape and manage departmental culture (Harris 2000).

The results of this study are consistent with findings by Thomas et al. (1997) who contend that departments are differentially effective which suggests that departments are at different levels of growth and therefore will require different improvement strategies and these strategies need to fit the 'growth- state' of the department. The departments can be identified as falling into three types as identified by Harris (2000). These three strategies were also reported by schools that have tackled within school variation in UK schools (National College for School Leadership 2004).

Ineffective Departments: They are stuck and need a high level of external support to provide knowledge about departmental improvement strategies. These departments need a number of early interventions and changes such as staff development which should focus on teaching and learning, building collaboration and changes in leadership. It has been shown that inappropriate leadership is a consistent feature of ineffective departments. Any collaboration and planning must be focused upon what happens in classrooms to improve student achievement and developing the potential of all staff. Time needs to be allocated for collaboration, developmental work, through observing each other and sharing of ideas.

Underachieving Departments: These must focus on teaching and learning as well as building the capacity of all staff to support these. There is a need for change in the leadership approaches and every effort made to harness the energy and optimism of new staff.

Moving Department: These must articulate and reinforce department values and expectations at every opportunity. It is important to solicit students' views on how their learning can be improved. This also encourages ownership for improvement by students.

The results on stability show that school effects may not be stable even over relatively short periods of time, for instance, three years considered in this study. Schools may be effective one year and not the other. The subjects were also not stable across the cohorts. This shows that the departmental effects can vary for relatively short periods of time. This is consistent with the study by Nuttall et al. (1989) who cautioned about any study of school effectiveness that relies on measures of outcome in just a single year or just a single cohort. They suggest that long time series are essential for a proper study of stability over time; three years being the minimum to identify a trend (Thomas et al. 1997).

Instability across years per subject will indicate a strong impact on individual teachers on student achievement (Luyten 1994). Consistent with the findings by Thomas et al. (1997), the correlation within subjects across the years were lower than the overall or total score by schools. This shows that school effects are more stable than departmental effects.

Schools which have developed a positive school culture and in which a tradition of high achievement has been acculturated may experience greater stability than others. The results on consistency and stability of school effects have important implications for bodies that report on the performance of schools, for example, the Botswana Examination Council. Value added measures give an accurate picture of a school performance. Effectiveness measures based on one outcome are likely to result in biased assessment of schools. A more appropriate indicator should be different criterion variables that will reveal the strong and weak areas of the school to eliminate complacency by the schools doing well in overall and also recognise the efforts by the schools that are not effective on overall results in some areas.

The results have implications for theory development. For school improvement, the department must be integrated as a unit of analysis. Harris (2001) argues that this integration moves the field a step closer to enquiry that there is greater linkage between process of school improvement and the reality of organizational change.

Achievement in a subject has a student effect as well. Further research is needed to investigate the attitudes of students towards these subjects. For schools that consistently do well, case study work is needed to document the factors that may be contributing to the levels of performance.

REFERENCES

- Blanchard J 2002. *Teaching and Targets: Self-evaluation and School Improvement*. London: Routledge Falmer.
- Brown M, Rutherford D, Boyle B 2000. Leadership for school improvement: The role of the Head of Department in UK secondary schools. *School Effectiveness and School Improvement*, 11(2): 232-258.
- Goldstein H 1999. Multilevel Statistical Models. From< <http://www.arnoldpublishers.com/support/goldstein.htm>> ((Retrieved on 23 October 2007).
- Gray J, Jesson D, Goldstein H, Hedger K, Rasbash 1995. A multi-level analysis of school improvement: Changes in school's performance. *School Effectiveness and School Improvement*, 6(2): 97-114.
- Doolaard S 2002. Stability and change in results of schooling. *British Educational Research Journal*, 28(6): 773-787.
- Harris A 2000. Effective leadership and departmental improvement. *International Journal of Research and Method in Education*, 23(1): 81-90.
- Harris A 2001. Department improvement and school improvement: A missing link? *British Educational Research Journal*, 27(4): 477-486.
- Highfield C 2010. Disparity in student achievement within and across secondary schools: An analysis of department results in English, maths and science in New Zealand. *School Leadership and Management*, 30(2): 171-190.
- Hopkins D 2001. *School Improvement for Real*. London: Routledge Falmer.
- Luyten H 1994. Stability of school effects in Dutch secondary education: The impact of variance across subjects and years. *International Journal of Educational Research*, 21(2): 197-216.
- Luyten H 1998. School effectiveness and student achievement, consistent across subjects? Evidence from Dutch elementary and secondary education. *Educational Research and Evaluation*, 4(4): 281-306.
- Luyten H 2003. The size of school effects compared to teacher effects: An overview of the research literature. *School Effectiveness and School Improvement*, 14(1): 31-51.
- Jakubowski M 2008. Implementing Value Added Models for School Assessment. EUI working papers RSCAS 2008/06. From< <http://www.eui.eu/RSCAS/Publications>> ((Retrieved on 28 April 2010).
- Kyriakides L 2004. Differential school effectiveness in relation to sex and social class: Some implications for policy evaluation. *Educational Research and Evaluation*, 10(2): 141- 161.
- Kyriakides L 2006. Introduction international studies on educational effectiveness. *Educational Research and Evaluation*, 12(6): 489-497.
- Kyriakides L, Creemers BPM 2008. A longitudinal study on the stability over time of school and teacher effects

- on student outcomes. *Oxford Review of Education*, 34(5): 521-545.
- Ma X 2000. Socioeconomic gaps in academic achievement within schools: Are they consistent across subject areas. *Educational Research and Evaluation*, 6(4): 337-355.
- Ma X 2001. Stability of school academic performance across subject areas. *Journal of Educational Measurement*, 38(1): 1-18.
- Mayston D 2007. Educational Value Added and Programme Evaluation. University of York. From <<http://www.dfes.gov.uk/research/data/uploadfiles/RW87.pdf>> (Retrieved on 20 June 2007).
- Mohiemang I 2009. *Effective Schools and Learners' Achievement in Botswana Secondary Schools: An Education Management Perspective*. D. Ed. Thesis, Unpublished. Pretoria: University of South Africa (Unisa).
- National College for School Leadership 2004. Tackling Within School Variation: A Leading Practice Seminar. From <<http://www.highreliability.co.uk>> (Retrieved on 25 May 2010).
- Nuttall D, Goldstein H, Prosser R, Rasbash J 1989. Differential school effectiveness. *International Journal of Educational Research*, 13(7): 769-776.
- Peng WU, Thomas S, Yang, X, Li J 2006. Developing school evaluation methods to improve the quality of schooling in China: A pilot 'value added' study. *Assessment in Education: Principles, Policy and Practice*, 13(2): 135-154.
- Rasbash J, Steele F, Browne W, Prosser B 2008. *A User's Guide to MLwiN Version 2.0.2*. Centre for Multilevel Modelling, University of Bristol.
- Sammons P 2006. The contribution of international studies on educational effectiveness: Current and future directions. *Educational Research and Evaluation*, 12(6): 583-593.
- Sammons P, Nuttall D, Cuttance P 1993. Differential school effectiveness: Results from a reanalysis of inner London Education Authority Junior School Project Data. *British Educational Research Journal*, 19: 381-405.
- Scheerens J, Bosker RJ 1997. *The Foundations of Educational Effectiveness*. Oxford: Pergamon Press.
- Teddlie C, Reynolds D, Sammons P 2000. The methodology and scientific properties of school effectiveness research. In: D Reynolds, C Teddlie (Eds.): *The International Handbook of School Effectiveness Research*. London: Falmer Press, pp. 53-133.
- Thomas S, Mortimore P 1996. Comparison of value added models for secondary school effectiveness. *Research Papers in Education*, 11(1): 5-33.
- Thomas S, Sammons P, Mortimore P, Smees R 1997. Stability and consistency in secondary schools' effects on students' GCSE outcomes over three years. *School Effectiveness and School Improvement*, 8(2): 169-197.