

IMPACT OF SCHOOL SELF-EVALUATION TRAINING ON IMPROVEMENT OF SELECTED DEPARTMENTS IN THE COLLEGE OF EDUCATION, UCC

Kenneth Asamoah-Gyimah
*University of Cape Coast,
Cape Coast,
Ghana.*

Abstract

This study examined the impact of school self-evaluation training on improvement of three Departments in the College of Education, UCC Ghana. All the lecturers at the three departments were sampled using the censors' survey. The participants from the three selected departments were assigned to one of the two training groups or the control group. The study employed the quasi experimental pro-test, post-test control group design. The research instruments used to generate data for the study were Personal Data Questionnaire (PDQ), and Institutional Self-Evaluation Questionnaire (ISEQ). Two hypotheses were formulated to guide the study. All the hypotheses were tested using the two-way analysis of covariance (ANCOVA) statistic at 0.05 level of significance. A post hoc test analysis was done using the Fisher's Least Square Method for all results found to be statistically significant. The results of the data indicated that all the two null hypotheses tested were rejected in favour of the treatment groups. The findings of the study revealed that establishment of School Self-Evaluation (SSE) mechanisms which are in line with the knowledge-base of Educational Effectiveness Research (EER) and the school stakeholders' involvement in defining the criteria of School Self-Evaluation (SSE) went a long way in boosting the departments' ability In evaluating themselves, however, the establishment of SSE mechanisms which are in line with the knowledge-base of EER had the strongest impact. The study has shown that the establishment of SSE mechanisms which are in line with the knowledge-base of EER significantly exhibit a higher performance in the structure of the curriculum and the overall quality of attainment indicators than those

exposed to school stakeholders' involvement in defining the criteria of SSE. These findings were placed within the existing body of knowledge and their implications for educational practices were discussed.

Keywords: School Self-Evaluation, School Self-Evaluation Mechanism, Knowledge-Base of Educational Effectiveness Research, School Stakeholders.

Improving the quality of education is currently a central concern of educational policy makers and implementers in many countries of which Ghana is no exception. Recent decades have witnessed a remarkable rise in the regulation of public services and servants, education being a case in point. According to Martin (2005), external evaluation and inspection have been important elements of this trend. He continued that increasingly, however, as the limitations of external surveillance systems have become clear, the concept of internal or self-evaluation has grown in importance. It is argued that enabling individual schools and lecturers to self-evaluate effectively is a complex task that will require help and support from the community of professional evaluators (McNamara and O'Hara, 2008a).

Contemporary educational systems are characterized by increased demands for effectiveness and quality as a result of the increased investment in education, and accountability demands by parents and society. It is argued that the above conditions require institutions to be involved in a continuous process of improvement. The rapidly growing interest in the fields of school effectiveness and school improvement is a response to these trends (Teddlie, 2010). The idea of schools undertaking a systematic process of self-evaluation seems at first sight a simple means of assessing the effectiveness of a school and finding ways of making it better. However, there is growing evidence that the process provides a valuable tool for teachers and schools to assess objectively how well they are doing ((Blok, Slegers, & Karsten, 2005); Moulten, 2009). Evaluation is the determination of a thing's value. In education, it is the formal determination of the quality, effectiveness or value of a programme, product, project, process and objective of curriculum (Worthen, Sanders & Fitzpatrick 2004).

School self-evaluation (SSE) concerns a type of educational evaluation at school level that is initiated and at least, partly controlled by the school itself. The Department of Education, Pretoria South Africa (2002), sees school self-evaluation as a school-based evaluation carried out by lecturers, heads of departments, departmental boards, Faculty and college governing boards, Academic boards and institutional

Communities at large. Through self-evaluation, institutions are able to prepare effectively for external evaluation and ultimately school self-improvement.

In the experience of many lecturers and students, school and classroom observations by visiting/external supervisors have failed to touch the real day-today experiences of students and their lecturers (Coleman, 1995). Coleman continues that in one secondary school in his study, students warned them to be wary of using visitors' impressions as a source of evidence. The students said they have become very well-trained on how to showcase the school at its best for outsiders and supervisors.

Improving the quality of education is currently a central concern to educational policy in many countries. Several countries have either created or are working on legislation and monitoring in the field of School Self-Evaluation (SSE), which stresses the schools' own responsibility for quality (Hofman, Hofman & Gray, 2010). Different approaches to establishing SSE mechanisms have been employed by various educational organizations as a means of achieving school effectiveness (Kyriades & Creemers 2008; Hofman, Hofman & Gray 2010).

The first approach to SSE is related to the assumption that the involvement of school stakeholders in defining the criteria of SSE may eventually encourage their active participation in using SSE for improvement purposes (Macbeath, 2004). Educational problems are typically complex, uncertain, multi-scale and affect multiple actors and agencies. This demands transparent decision-making that is flexible to changing circumstances, and embraces a diversity of knowledge and values. To achieve this, stakeholder participation is increasingly being sought and embedded into school-level decision-making processes, from local to international scales (Richards, Blackstock & Carter 2004).

The second approach is concerned with the establishment of a climate in the school that supports change. This approach is not only based on findings of school improvement projects but also on the view of schools as mini political systems with diverse constituencies. The terminology micro politics of education emerged in clearly articulated form in the research literature within the past 30 years (Hoyle & Skria, 1999). Micro politics recognize divergence of interest, multiple sources of power, and the potential of conflict within the school setting. Such a lens allows for the possibility that coalitions and conflicts may occur both across and within organizations such as schools (Firestone & Fisler, 2002).

The third approach is based on the assumption that the knowledge-base of Educational Effectiveness Research (HER) should be taken into account in developing SSE mechanisms (Teddle & Stringfield, 2007). A major element of this approach is the emphasis on the evidence stemming from theory and research. Thus,

the value of a theory-driven approach is stressed. For example, the need to collect multiple data about student achievement, the classroom and school processes is emphasized by making use of a theoretical framework based on the main findings of EER. A distinctive feature of EER is that it does not only refer to factors that are important for explaining variations in educational effectiveness but also attempts to explain why these factors are important by integrating different theoretical conditions to effectiveness (Heck & Moriyama 2010).

Quality Assurance has become a major issue in contemporary university administration worldwide. This has been occasioned by a number of reasons including competitiveness on the job market which favour high quality graduates. This enjoins universities to ensure the quality of their academic programmes and delivery mechanisms. Furthermore the multiplicity of higher education institutions in recent times and the emergence of social accountability require managers of universities to constantly improve quality and promote transparency in order to safeguard public interest and confidence in their awards. Finally the rapid expansion in enrollment in the face of dwindling resources in many African universities, imposes a burden with the potential of compromising quality. The Association of African Universities has accordingly obliged member institutions to set up quality assurance units and harmonize activities at national, sub-regional and regional levels. Such a unit known as the Directorate of Academic Planning and quality assurance has been established in the University of Cape Coast since 2001. The directorate has collaborated with its counterparts in sister Ghanaian Universities and made reasonable impact on the delivery of core business of the university. There is now the need to put in place mechanisms at the university to facilitate the realization of institutional vision at the college and departmental levels.

Research Hypotheses

1. There is no significant difference in post training scores on the structure of the curriculum indicator among respondents exposed to the knowledge-based Educational Effective Research (EER), School Stakeholders' Involvement in developing SSE mechanisms and the control group.
2. There is no significant difference in post training scores on the overall quality of attainment indicator among respondents exposed to the knowledge-based Educational Effective Research (EER), School Stakeholders' Involvement in developing SSE mechanisms and the control group.

Methodology

The study employed the quasi-experimental pretest-posttest control group design to explore the effects of training and the relationship that exists among the

training groups and the control group. Eighty-eight lecturers and 129 students were sampled from three departments from the college of education for the study. The participants from the three selected departments were assigned to one of the two training groups or the control group. The departments, lecturers and third year students were selected using the multistage sampling technique. The sampling techniques include the simple random, cluster and purposive. The instruments used to obtain the relevant data for the study was the Institutional Self-Evaluation Questionnaire (ISEQ) The school self-evaluation indicators: structure of the curriculum; overall quality of attainment; teaching and learning process; were the independent variables.

Results

All the two hypotheses were tested using two-way analysis of covariance (ANCOVA) statistics. All the hypotheses were tested at 0.05 level of significance.

Hypothesis One: There is no significant difference in post treatment scores on the structure of the curriculum indicator among respondents exposed to the knowledge-base of Educational Effective Research (EER), School Stakeholders' Involvement in developing SSE mechanisms and the control group. The hypothesis was tested using the two-way analysis of covariance (ANCOVA) statistics. The result of the analysis is presented in Tables 1, 2 and 3.

Table 1: Descriptive Data on the Structure of the Curriculum Indicator among Respondents across Experimental Conditions and Status

Group	Status	N	Pre Test		Post Test		Mean Diff.
			Mean	SD	Mean	SD	
Knowledge-base of EER	Teacher	29	13.93	2.48	25.31	2.22	-11.38
	Student	43	13.72	2.75	23.60	2.39	-9.88
	Total	72	13.81	2.63	24.29	2.46	-10.48
School Stakeholders' involvement	Teacher	31	18.97	2.95	24.06	2.41	-5.09
	Student	45	14.58	2.57	22.22	1.98	-7.64
	Total	76	16.37	3.48	22.97	2.33	-6.60
Control	Teacher	28	16.46	2.89	16.96	3.14	-0.50
	Student	41	16.88	2.11	17.73	2.21	-0.85
	Total	69	16.71	2.44	16.23	2.68	0.45
Total	Teacher	88	16.51	3.45	22.22	4.47	-5.71
	Student	129	15.02	2.81	20.62	4.04	-5.60
	Total	217	15.63	3.16	21.27	4.28	-5.64

From Table 1, participants exposed to knowledge-base of EER had the highest mean difference of 10.48 followed by those exposed to school stakeholders

involvement with mean difference of 6.60. The control group had the lowest mean difference of 0.45, on the structure of the curriculum indicator.

To determine whether significant difference in the structure of the curriculum existed across the experimental conditions, the two-way analysis of covariance (ANCOVA) statistics was used to analyze the data. The result of the analysis is presented in Table 2.

Table 2: Two-Way Analysis of Covariance (ANCOVA) on Difference in Scores the Structure of the Curriculum across Experimental Conditions

Source	Sum of Square	Degree of Freedom	Mean of Squares	F-Value
Corrected Model	2767.35	6	461.23	81.59*
Covariate -	0.01	1	0.01	1.00
Exp. Group	2381.92	2	1190.96	210.67*
Status	123.85	1	123.85	21.91*
Exp. Group vs. Status	3.22	2	1.61	0.29
Error	1187.15	210	5.65	
Total	6463.50			

* Significant at 0.05; df = 1, 2, & 210; Critical F = 3.89 & 3.04 respectively

The results in Table 2 show that a calculated F-value of 210.67 resulted as the difference in the structure of the curriculum indicator among the three experimental groups due to experimental conditions. This calculated F-value of 210.67 is statistically significant since it is greater than the critical F-value of 3.04 given 2 and 210 degrees of freedom at 0.05 level of significance. This implies that there is significant difference in the posttest scores on the structure of the curriculum indicator among the three groups. Table 2 also shows that a calculated F-value of 21.91 resulted as the influence of status of participants on the structure of the curriculum. This calculated F-value is significant since it is greater than the critical F-value of 3.89 given 1 and 210 degrees of freedom at 0.05 level of significance. This implies that there is significant difference between lecturers and students in the post test scores on the structure of the curriculum indicator. Lecturers generally had higher mean scores than students on this indicator. The results in Table 2 show that the total mean for lecturers (Mean = 22.22, SD = 4.47) is significantly higher than that of students (Mean = 20.62, SD = 4.04) on the structure of the curriculum indicator.

Further evidence from Table 2 revealed that the interaction effect between experimental condition and participants' status (lecturer or student) on the structure of the curriculum indicator resulted in F-value of 0.29 which is not significant since it

is less than the critical F-value of 3.04 given 2 and 210 degrees of freedom at 0.05 level of significance.

Due to the significant effects of the experimental conditions on the structure of the curriculum indicator, post hoc test comparisons were carried out to determine which pair of the group means exhibited significant difference as well as the trend of the difference using Fisher’s protected t-test is reported in Table 3.

Table 3: Fisher’s Protected t-test on Differences in Participants Performance on the Structure of the Curriculum Indicator

Group	Knowledge-Base of EER n = 72	School Stakeholders’ Involvement n = 76	Control n=69
Knowledge-base of EER	24.29 ^a	3.37*	20.10*
School Stakeholders’ involvement	1.32	22.97	17.02*
Control	8.06	6.74	16.23

a = group means are in the diagonal; difference in group means are below the diagonal while protected t-values are above the diagonal

* Significant at 0.05

The data in Table 3 show that participants exposed to the knowledge-base of EER significantly performed higher in the structure of the curriculum indicator than those exposed to School Stakeholders’ Involvement in School Self-evaluation (cal t=3.37; df=146; critical t=L96; p<0.05). Again, participants exposed to the knowledge-base of EER significantly performed better on the structure of the curriculum indicator than the control group (cal t=20.10; df = 139; critical t = 1.96; p<0.05). Similarly, participants treated with School Stakeholders’ Involvement in school self-evaluation significantly have higher performance on the structure of the curriculum indicator than the control group (cal t-17.02; df=145; critical t=1.96). The null hypothesis is, therefore, rejected.

Hypothesis Two: There is no significant difference in post treatment scores on the overall quality of attainment indicator among respondents exposed to the knowledge-base of EER, School Stakeholders Involvement in developing SSE mechanisms and the control group. The result of the analysis using the two-way analysis of covariance (ANCOVA) is presented in Tables 4, 5 and 6.

Table 4: Descriptive Data on the Overall Quality of Attainment Indicator among Respondents across Experimental Conditions and Status

Group	Status	N	Pre Test		Post Test		Mean Diff.
			Mean	SD	Mean	SD	
Knowledge-base of EER	Teacher	29	12.90	1.70	19.62	1.12	-6.72
	Student	43	11.33	1.54	16.47	1.56	-5.14
	Total	72	11.96	1.77	17.74	2.09	-5.78
School Stakeholders' involvement	Teacher	31	12.90	1.96	17.87	1.65	-4.97
	Student	45	11.33	1.90	15.53	1.78	-4.20
	Total	76	11.76	2.14	6.49	2.07	-4.73
Control	Teacher	28	12.61	1.03	12.57	1.45	0.04
	Student	41	11.27	1.76	9.90	2.27	1.37
	Total	69	11.81	1.64	10.99	2.37	0.82
Total	Teacher	88	12.81	1.61	16.76	3.29	-3.95
	Student	129	11.19	1.74	14.05	3.43	-2.86
	Total	217	11.84	0.86	15.15	3.62	-3.31

From Table 4, participants exposed to knowledge-base of EER had the highest mean deviation score of 5.78 followed by those exposed to school stakeholders' involvement with mean difference of 4.73 while the control group had the lowest mean difference of 0.82. To determine whether significant differences in the overall quality of attainment indicator exist, the two-way ANCOVA statistics was used. The result is presented in Table 5.

Table 5: Two-Way Analysis of Covariance (ANCOVA) on Difference in scores on the Overall Quality of Attainment Indicator across Experimental Conditions

Source	Sum of Squares	Degree of Freedom	Mean of Squares	F-value
Corrected Model	2205.67	6	367.61	124.45*
Covariate	0.36	1	0.36	0.12
Exp. Group	1758.54	2	879.27	297.67*
Status	306.84	1	306.84	103.87*
Exp. Group vs. Status	6.12	2	3.06	1.04
Error	620.31	210	2.95	
Total	4897.84			

* Significant at 0.05; df = 1, 2, & 210; Critical F = 3.89 & 3.04 respectively.

The results in Table 5 show that a calculated F-value of 297.67 resulted as the difference in overall quality of attainment indicator among the three groups due to experimental conditions. This calculated F-value of 297.67 is statistically significant since it is greater than the critical F-value of 3.04 given 2 and 210 degrees of freedom at 0.05 level of significance. This implies that there is significant difference in the posttest scores on the overall quality of attainment indicator among the three groups. Table 5 also shows that a calculated F-value of 103.87 resulted as the influence of status of participants on the overall quality of attainment indicator. This calculated F-value is significant since it is greater than the critical F-value of 3.89 given 1 and 210 degrees of freedom at 0.05 level of significance. This implies that there is significant difference between lecturers and students in the post test scores on the overall quality of attainment indicator. Lecturers generally had higher mean scores than students on this indicator. The results in Table 5 show that the total mean for lecturers (Mean = 16.76, SD = 3.29) is higher than that of students (Mean = 14.05, SD = 3.43) on the overall quality of attainment indicator. Further evidence from Table 5 revealed that the interaction effect between experimental condition and participants' status on the overall quality of attainment indicator resulted in calculated F-value of 1.04 which is not significant since it is less than the critical F-value of 3.04 given 2 and 210 degrees of freedom at 0.05 level of significance.

Due to the significant effects of the experimental conditions on the overall quality of attainment indicator, the post hoc test comparisons were carried out to determine which pair of the group means exhibited significant difference as well as the trend of the difference using Fisher's protected t-test reported in Table 6.

Table 6: Fisher's Protected t-test on Differences in Participants Performance on the Overall Quality of Attainment Indicator

Group	Knowledge-Base of EER n = 72	School Stakeholders' Involvement n = 76	Control n=69
Knowledge-base of EER	17.74 ^a	4.42*	23.30*
School Stakeholders' involvement	1.25	16.49	19.22*
Control	6.75	5.50	10.99

a = group means are in the diagonal; difference in group means are below the diagonal while protected t-values are above the diagonal

* Significant at 0.05

The data in Table 6 show that participants exposed to the knowledge-base of EER significantly demonstrated a higher performance in the overall quality of attainment Indicator than those exposed to School Stakeholders' Involvement in School Self-evaluation (cal t=4.42; df=146; critical t=1.96; p<0.05). Again,

~~participants~~ exposed to the knowledge-base of EER significantly showed a higher performance than the control group (cal $t=23.30$; $df=139$; critical $t=1.96$; $p<0.05$). Similarly, participants exposed to School Stakeholders' Involvement in school self-evaluation significantly have higher performance in the overall quality of attainment indicator than the control group (cal $t=19.22$; $df=143$; critical $t=1.96$). The null hypothesis is, therefore, rejected.

Discussion of Findings

The first hypothesis sought to find if a significant difference exists on participants' scores on the structure of the curriculum across experimental conditions. The two-way ANCOVA together with the descriptive statistics was used in analyzing the data for this hypothesis. The results in Tables 1, 2 and 3 revealed that there is significant difference in the posttest scores on the structure of the curriculum indicator among the participants exposed to the two different approaches to school self evaluation and the control group. The post hoc analysis, using Fisher's Protected t-test, also revealed that participants exposed to knowledge-base of EER significantly performed higher in the structure of the curriculum indicator than those exposed to school stakeholders' involvement. The two experimental groups also significantly exhibited higher performance in the structure of the curriculum indicator than the control group. Also lecturer-participants significantly demonstrated higher performance in the structure of the curriculum indicator than the students who participated in the study. This was reflected in their post test total means (lecturers Mean = 22.22, SD = 4.47 and students Mean = 20.62, SD = 4.04). The null hypothesis was, therefore, rejected.

The better performance among participants exposed to the two treatments was expected since the participants had been provided with guidelines on curriculum assessment with emphasis on the breadth and balance across elements, timetabling and arrangements for pupil choice as well as courses and programmes. This finding is in line with the findings of Lockheed and Levin (1993) who argue that creating effective schools in developing countries requires necessary inputs in terms of curriculum.

The second finding of the study is that participants in the two experimental groups (those exposed to establishment of SSE mechanisms which are in line with the knowledge-base of EER and school stakeholders' involvement in defining the criteria of SSE) significantly demonstrated higher performance on the Overall Quality of Attainment Indicator than those in the control group. In arriving at this finding, the two-way ANCOVA was used. A further analysis, using Fisher's Protected T-test also showed that participants exposed to the establishment of SSE mechanisms which are in line with the knowledge-base of EER demonstrated higher performance in the overall quality of attainment indicator than those exposed to school stakeholders'

involvement in defining the criteria of SSE. Also participants exposed to school stakeholders' involvement in defining the criteria for SSE exhibited higher performance in the overall quality of attainment indicator than those in the control group. The findings also revealed that participants who were teachers significantly demonstrated higher performance in the overall quality of attainment indicator than the students who participated in the study. This was reflected in their post test total means (teachers Mean = 16.76, SD = 3.29 and students Mean = 14.05, SD = 3.43). This finding is not surprising and is in the expected direction since, in Ghana, lecturers are more involved in determining the overall quality of students' attainment through the various formative and summative evaluation practices like the Continuous Assessment System and the end-of-Semester examination analysis. The null hypothesis was, therefore, rejected.

This finding is in agreement with the works of Kellaghan and Greaney (2004); Pennycook (1993) Velez, Schiefelbein & Valenzuela (1993) which all revealed that assessing and monitoring student academic progress/attainment is considered an important factor in promoting school effectiveness in both developed countries and developing countries. From the perspectives of developing countries, Kellaghan and Graney (2004) have been the key proponents of singling out the use of the overall quality of attainment and examination reforms to engineer change at the levels of educational policy to enhance quality through national examinations.

Conclusions

Beyond the fact that all two experimental groups (establishing SSE mechanisms which are in line with the knowledge-base of EER and the involvement of school stakeholders in defining the criteria of SSE) had better results than the control group, implying that SSE can contribute in establishing effective school improvement strategies, the establishment of SSE mechanisms which are in line with the knowledge-base of EER had the strongest impact. Establishing SSE mechanisms which are in line with the knowledge-base of EER is thus, the most effective way of establishing SSE on the structure of the curriculum, overall quality of attainment, support for pupils, ethos and the resources. The establishment of SSE mechanisms which are in line with the knowledge-base of EER group did not differ significantly from those exposed to the involvement of school stakeholders in defining the criteria SSE on the teaching and learning processes as well as the management leadership and quality assurance indicators.

Through the findings it can be concluded that if the university authorities - Lecturers, Heads of departments, Deans, Pro-vests, Pro-vice Chancellors and the Vice chancellor - can present an enabling atmosphere that will make it possible for the University - Colleges, Faculties, departments and section/units - to apply self-evaluation in the university setting systematically, it will go a long way to sustain the

best practices of the numerous interventions and reforms that the universities in Ghana have benefitted. It will also go a long way to create the sense of ownership, as a result of participation, which leads to desired forms of commitment and motivation of all the stakeholders in the school.

Implications for Educational Practice

The issue of school evaluation is becoming more urgent and ever more complex issue in education that has challenged school administrators, lecturers and students. Lecturers and other university staff in most cases argue that evaluation schemes, especially external evaluation, constrain their autonomy. They, therefore, oppose imposed evaluation schemes, especially when sanctions are at stake. The findings of this study provide empirical answers to some of the issues/challenges that stakeholders in Ghanaian Educational system have been encountering. Several educational implications could be derived from this study to improve educational practice in Ghanaian universities.

Firstly, the fact that all the two experimental groups had better results implies the two approaches to SSE (establishing SSE mechanisms which are in line with the knowledge-base of EER and the involvement of school stakeholders in defining the criteria of SSE) are effective. Universities in Ghana can, therefore, adopt these approaches to SSE in the schools as a means of ensuring quality in the country's educational provision. It will be more expedient to orientate staff at the universities and students to these approaches through in-service training programmes than to rely solely on external supervisors and others who are not regular and insiders in many cases.

References

- Blok, H., Slegers, P. & Karsten, S. (2005). *School Self Evaluation in Primary Education*. Amsterdam: SCO-Kohnstamm Instituut.
- Coleman, J. S. (1995). Families and Schools. In Coleman, J. S. *Equality and Achievement in Education*. (pp. 325 – 340) Boulder, CO: Westview Press.
- Firestone, W. A. & Fisler, J. L. (2002). Politics, Community and Leadership in a School University Partnership. *Educational Administration Quarterly*, 38 (4), 449 – 493.
- Heck, R. H. & Moriyama, K. (2010). Examining Relationships among Elementary Schools' Contexts, Leadership, Instructional Practices, and Added-Year Outcomes: A Regression Discontinuity Approach. *School Effectiveness and School Improvement*. 21 (4), 377 – 408.

- Hofman, R. H., Hofman, W.H., & Gray, J. M. (2010). Institutional Contexts and International Performances in Schooling: Comparing Patterns and Trends over Time in International Surveys. *European Journal of Education*, 1, 153 – 173.
- Hoyle, R. J. & Skria, L. (1999). The Politics of Superintendent Evaluation. *Journal of Personnel Evaluation in Education*, 13 (4), 405 – 519.
- Kellaghan, T. & Greaney, V. (2004). *Assessing Student Learning in Africa*. Washington, D.C.: World Bank.
- Kyriakides, L. & Creemers, B. P. M. (2008). Using a Multi-Dimensional Approach to Measure the Impact of Classroom-Level Factors upon Student Achievement: A Study Testing the Validity of the Dynamic Model. *School Effectiveness and School Improvement*, 19, 183-205.
- Lockheed, M. E. & Levin, H. M. (1993). Creating Effective Schools. In H. M. Levin & M. E. Lockheed (Eds.), *Effective Schools in Developing Countries* (1- 19). London: The Falmer Press (with the World Bank).
- MacBeath, J. (2004). *The Self-Evaluation File*. Glasgow: Learning Files Scotland.
- McNamara, G. & O'Hara, J. (2008a). The Importance of the Concept of Self-Evaluation in the Changing Landscape of Education Policy. *Studies in Educational Evaluation*, 34, 173 – 179.
- Martin, S. (2005). Evaluation, Inspection and the Improvement Agenda: Contrasting Fortunes in an Era of Evidence-Based Policy Making. *Evaluation*, 11 (4), 496-504.
- Moulten, J. (2009). Improving the Quality of Education: What has the World Bank Learned. *Journal of Dental Education*, 73, 883-887.
- Pennycuik, D. (1993). *School Effectiveness in Developing Countries. A Summary of the Research Evidence*. Great Britain: Overseas Development Administration.
- Richards, C., Blackstock, K. L., & Carter, C. E. (2004). *Practical Approaches to Participation*. Aberdeen: SRG Policy Brief No. 1 Macauley Land Use Research Institute.
- Teddlie, C. (2010). The Legacy of the School Effectiveness Research Tradition. In A. Haegreaves, A. Lieberman, M. Fullan & D. Hopkins (Eds.). *The Second International Handbook of Educational Change*. Dordrecht: Springer.

- Teddlie, C. & Stringfield, S. (2007). A History of School Effectiveness and Improvement Research in the USA Focusing on the Past Quarter Century. In T. Townsend (Ed.), *International Handbook of School Effectiveness and Improvement* 131-166. Dordrecht, Netherlands: Springer.
- Velez, E., Schiefelbein, E., & Valenzuela, J. (1993). *Factors Affecting Achievement in Primary Education: A Review of the Literature for Latin America and the Caribbean*. Washington, D.C.: Department of Human Resources Development and Operations Policy, World Bank.
- Worthen B. R., Sanders, J. R., & Fitzpatrick, J. L. (2004). *Educational Evaluation: Alternative Approaches and Practical Guidelines*. (3rd ed.). Boston: Allyn & Bacon.