EUROPEAN JOURNAL OF MULTIDISCIPLINARY SCIENCES
EjMS
ISSN: 2421-8251

European Journal of Multidisiplinary Sciences (EjMS)
Volume 1, Issue 1 (e-ISSN: 2421-8251)

THE CONSISTENCY AND STABILITY OF SCHOOL EFFECT IN COLOMBIAN HIGH SCHOOL EDUCATION

Olga Rosalba Rodríguez-Jiménez[a]*

[a] *Universidad Nacional de Colombia, Ciudad Universitaria Edificio 212- of 225, Bogotá, Colombia*

Abstract

Consistency and stability are considered scientific properties of school effect. These properties are important in determining if school effect is constant along subjects and are perdurable during the time. The purpose of this research is to estimate continuity and stability in Colombian high school education during seven years over the past decade for the four main subjects in Colombia: mathematics, language, natural and social science. With this purpose in mind, a retrospective ex post facto study was conducted. The analysis was done for 7, 626 schools and multilevel models of three levels and longitudinal multilevel models were used. As adjusted variables, socio-economic and sociocultural levels and gender of students were used. The school variable was the tuition. The results indicated that the correlation was over 0.39 for all the subjects; the higher correlations were found among language, natural and social science and the school effect showed change along the time for every subject and by the tuition. The Colombian educational system displays consistency within subjects, but its school effect is not stable and is dependent on the socio-economic variable. The paper discusses the implications of these results for the Colombian educational system and the effect for quality and equity, and their relationship with the development of the country.

Keywords: School effect, Colombian high school, continuity and stability, school effectiveness

*Corresponding author.
E-mail address: orrodriguezj@unal.edu.co

doi: 10.15405/ejms(2421-8251).2016.1.2

1. Introduction

The study of the scientific properties of the school effect is a research field that has developed in line with the growing interest in educational quality. The four principal scientific properties are consistency, stability, differential effectiveness and continuity. Consistency refers to the congruency among different subjects in the curriculum and the resulting cognitive and socio-affective achievement (Murillo, 2005a). Stability refers to the school effect permanency over time. Differential effectiveness allows researchers to

study if the school effectiveness is constant for students of different ethnicity and sociocultural origin. Finally, continuity refers to the school effect permanency over a duration of time after students finish their studies.

Consistency has been studied a lot; the majority of the studies have been conducted in secondary education between mathematics and language. The methods for these studies vary; some use correlations, time series, discontinuity of regression, growth curves, survival models, and the most frequent, multilevel models. The results show a low consistency and changes at the school are not perdurable over time. Research on this topic has been done in Europe and USA but few in Latin America with particular focus on the associated factors (Rodríguez-Jiménez & Murillo, 2011), the diversity of the theory (Murillo, 2007b) and methodology. Because of this, it is difficult to build a strong research base and develop on it. Such studies are useful as their findings focus on the performance of educational system (Kyriakides, 2004) as reflected in the educational standards (Postlethwaite, 1994), accountability and the social responsibility of the education system.

Hence, in order to obtain information about the efficacy of the Colombian educational system in relation to its consistency and stability as well as its relationship with equality and its impact on the development of the country, this study was conducted to estimate the continuity and stability in Colombian high school education over a seven year period covering the 2000 decade in four main subjects in Colombia: mathematics, language, natural and social science.

2. Background on scientific properties associated with school effect

2.1. Consistency

This property is studied through the correlation among the residuals of the adjusted models (Murillo & Hernández-Castilla, 2011) with information gotten at the same time (Thomas et al., 1997) and different variables of criteria (Kyriakides & Creemers, 2008). This property is related to the dimensionality of school effect (Kyriakides & Creemers, 2008). Studies on this property have been conducted in primary and preparatory levels using cognitive and socio-affective measures.

In elementary education, the most important study is by Mandeville and Anderson (1987) who found for 2,083 students of 423 schools in the USA, correlations between mathematics and language were between 0.63 to 0.70; the correlations among degrees were between 0.08 to 0.16. The results suggest that the school effect is consistent but remains at the same level. Another important study was done by Mortimore, Sammons, Stoll, Lewis, and Ecob (1988) who reported correlations between mathematics and reading at 0.41. The conclusion was that good results in one topic were related to the good results in others. Later, Sammons, Nuttall, and Cuttance (1993) did a re-analysis of this information and reported higher correlations (r=0.61 between mathematics and reading). In general, the results showed the correlations for the subjects is between 0.70 to 0.75 and for different criterial measures of 0.00 and 0.05 (Bosker & Scheerens, 1989).

In secondary education, Willms and Raudenbush (1989) in their longitudinal studies reported for Type A school effect a correlation between 0.46 to 0.73 and for Type B the effects were between 0.19 to 0.71 which according to them proves consistency among subjects. Similarly, Bosker and Scheerens (1989) found that the correlation among subjects was between 0.45 to 0.75 and between 0.35 to 0.70 among different criterial variables. These results suggest that if the school is successful in one subject, then it may

be successful in the others (Willms & Raudenbush, 1989). The magnitude of the correlations, lower than primary, could be effect for teacher and better measures in this level.

Thomas et al. (1997) found correlations of 0.25 (English literature and French) and 0.75 (total scores and English) in longitudinal studies done from 1990 until 1992 using multilevel models. The authors concluded that there are substantial differences among subjects for the schools and they posited that these results are due to the teachers. Besides, Cuttance (1987), Thomas and Mortimore (1996); Thomas and Goldstein (1994) and Fitz-Gibbon, Tymms, and Hazlewood (1990) found moderate correlations between English and Arithmetic with values between 0.46 to 0.48. These results contradict the ones obtained by Goldstein and Rasbash (1993) who found a correlation of 0.29 between Mathematics and English. Hence, they concluded that the ranking of schools depends on the adjusted variables and the curriculum measures more than the consistency among subjects.

In relation to the correlations for different product measures, Rutter et al.'s research (1979) reported positive correlations with values between 0.23 to 0.65 among subjects and between academic and social success and r=0.80 for class attendance and delinquency. Similarly, Mortimore et al. (1988) reported correlations for writing and class attendance, and Mathematics and self-concept. Despite these positive results, these researchers claim that cognitive and affective measures are independent. Opdenakker and Van Damme (2000) and Thomas (2001) ensure that the school effect is higher for cognitive compared to affective results. This may be because the latter has the least attention in the curriculum and their measure is imprecise (Reynolds et al., 2011).

In contrast, Iberoamerican research has reported consistency among subjects. Murillo (2007a) found moderate consistency with values nearest to 0.5 for Mathematics and language and correlations between cognitive and affective products measure (r= 0.26 and r=0.40 for school satisfaction at Mathematics and language respectively). For the same subjects, in Mexico, Zorrilla (2008) reported high consistency (r=0.86 in 2002 and r= 0.87 in 2003) and in Argentina, Cervini (2010) presented high correlations (r=0.78 in primary r= 0.80).

According to Murillo (2005a), the research in general reported moderate resistance with correlations close to 0.5. For this reason, it is important to include more subjects and different results variables. In relation to the products measure by Reynolds et al. (2011) acceptable consistency between subjects was found, with higher magnitude in primary than secondary schools. However, results are not conclusive about the relationship between cognitive and affective measures because there is little research on it.

2.2. Stability

This property is related to permanence by the time (Murillo, 2005a; Reynolds et al., 2011). This property refers to the characteristics of the school and teachers. To investigate stability, the correlation within residuals of models in two different periods of time (Scheerens & Bosker, 1997; Willms & Raudenbush, 1989) is studied. Due to this, in their analysis, researchers used longitudinal models and repeated measures. With regard to this property, authors such as Kyriakides and Creemers (2008) argued that it is necessary to study the theoretical approaches to find out about the source of the stability, for example, if the change is natural or programmed; whether it depends on school characteristics, teachers, or internal or state politics. In this sense, the authors ensure that the school effectiveness should be studied as a dynamic system because there is interaction between different factors such as students' background,

school characteristics and the educational system. In consequence, the schools can identify their weaknesses and take the necessary action to improve and increase their school effect (Kyriakides and Creemers, 2010).

A pair of studies in elementary education was conducted by Mandeville and Anderson (1987) and by Mandeville (1988). The first study showed correlations between 0.52 to 0.33 for different cohorts. This instability may be explained by the early age of students and their development. As for consistency, Bosker and Scheerens' (1989) study found correlations between 0.35 to 0.65 for different cohorts (years) and correlations between 0.10 to 0.65.

In secondary education, Bosker et al. (1988, as cited in Bosker & Scheerens, 1989) reported correlations between 0.40 to 0.80 for secondary levels. However, they claimed that it was over estimated because there is dependence inside the information as the measure was similar for students in the different levels and the criterial variable is accumulative along the time. Thomas et al. (1997) reported index of correlations to subjects between 0.38 to 0.92. The subjects with lowest stability were French and the highest was history. In general, they found higher stability to the total scores.

In USA, Willms and Raudenbush (1989) reported for the A school effects correlations of 0.87 to the total scores for 1980 and 1984; r=0.13 for Mathematics and r= 0.73 for English. In addition, for B school effect the correlation index was equal to 0.70 in the total score; r= -0.013 for Mathematics and r= 0.79 for English. Considering that these results reflected the school interaction over time, the authors concluded that the instability of the school effect was due to the subjects because of the change of teachers and the implementation of major or short-term education reforms. Bosker and Scheerens (1989) showed the results for cohorts between 0.70 to 0.95 and for levels between 0.25 to 0.90. They argued that the study stability within levels and signatures is an important topic (Luyten, 2004) because it can show the specificity and effectiveness of teaching and when these are high it may be due to the accumulation of the effects every year (Goldstein, 1997). If these are low, it could be due to the changes that the schools undergo in order to adjust to external circumstances.

In England, Mangan, Pugh, and Gray (2005) studied 541 secondary schools for a period of ten years (1992-2001); the variable was the score in GCSE. They used time series and did trajectories on the time frame. The results showed little increase; between 0.06 to 0.0124, where the rate of improvement was between 1.2 to 5.9. According to the geographic distribution, the schools displayed different patterns of improvement which were unstable. 63% of the schools posted a decrease in three or more years in the number of highest scoring students. The previous performance had a positive relationship with the current performance, but this effect reduces quickly with dynamic models for one previous year showed a value of 0.52 and for two previous years the value was 0.21. In relation to the prediction, the gender of the students and the management of the school were included as variables which resulted in 95% of the schools posting important differences for three years. The authors concluded that stability only occurs for a short time, maximum for three years and because of that, it was difficult to make predictions.

In Iberomerica, there is not much information about stability. Hence, it is necessary to do more research with different levels and background variables and criterial variables (Murillo, 2005a). In addition, the effectiveness and change of teachers should also be studied (Kyriakides & Tsangaridou, 2008).

3. Research design and methods

The current study is an ex post facto retrospective design which uses secondary exploration of data. The information comes from Instituto Colombiano para la Evaluación de la Educación -ICFES which is a government entity for monitoring and certification of quality education in Colombia.

3.1. Research questions

The questions that guide the study are as follows:

a) Does school effect change within subjects for the Colombian educational system?

b) Is school effect perdurable over time in the Colombian educational system?

Bearing in mind this purpose, the research investigated the estimation of consistency in high school for mathematics, language, natural and social science over seven years in the 2000 decade as well as the estimation of the stability of school effect for high school during the same time.

3.2. Participants

The sample is found in Table 1. The information shows that the percentage of girls is higher than boys for all years except 2000. To test for stability, the same schools were studied for the seven years; 6,844 schools and 49,210 students.

Table 1. Sample of the scientific estimation properties of school effect in high school

	Nº			%	
Year	States	Schools	Students	Girls	Boys
2000	33	7.626	314.560	45,5	54,5
2001	33	5.705	237.225	54,9	45,1
2002	33	9.704	315.486	54	46
2003	33	8.375	304.199	54	46
2008	33	7.742	379.429	54,1	45,9
2009	33	8.022	402.664	54,2	45,8
2010	33	10.977	441.318	54,2	45,8
Total			2.394.881		

3.3. Variables, test and questionnaire

The research included two variables: results and adjustment similar to other similar studies (Cervini, 2010; Murillo, 2005b, 2007a, 2008; OREALC/UNESCO, 2010; Zorrilla, 2008).

The product variable was scored using a standarized test for mathematics, language, natural science (physics, chemistry and biology) and social sciences (geography, history and philosophy). These results are scaled with average 50 and standard deviation 10; minimum 0 and maximum 100 (ICFES, 2011c).

The adjusted variables are for the students and schools. The variables for students were:

a. Gender (GENERO): 0= boys; 1= girls.
b. Mother tongue (LENGUA): 1= native language, 0= Spanish.
c. Sociocultural and socioeconomic indexes (ZNSC1, ZNE).

https://doi.org/10.15405/ejms(2421-8251).2016.1.2
eISSN: 2421-8251 / Corresponding Author: Olga Rosalba Rodríguez-Jiménez
Selection & Peer-review under responsibility of the Editors

The index was synthetic in accordance with the methodology of Becerra-Avella (2010). These are standarized variables with average 0 and standard deviation 1; these included parents' occupation and education and monthly income.

The variables for the school were:

a. Socio-economic level (NSE_ESCUELA), standarized with average 0 and standard deviation
b. Tuition (Pension): ordinal variable with seven categories.

The information about the variables was obtained through a standardized test and a 90-item questionnaire on students, parents, and schools' background. Until 2007, the standardized test contained 35 items covering each subject and this number was later reduced to 24 items (ICFES, 2005). The psychometric quality, done with Rasch model showed a good *infit* and *outfit* adjustment. The test had an equating process until 2000 (ICFES, mail in 2013). When the students registered for the test, they were given the questionnaire to fill up.

3.4. Data analysis

The analysis was done with MLWin with different phases. The first was the preparation of information which included depuration and preparing for the program. According to this process, 10% and 20% was eliminated; the main criteria was missing values and schools with less than ten students. The second was the estimation of every model without adjusted variables. In addition, the model was adjusted for maximum probability. Finally, the deviance was obtained to know the adjustment of models.

The consistency was calculated for mathematics, language, social and natural sciences with multilevel models of three levels: student, school and state with MLWin 2.3. Once the models were adjusted, the standardized residuals and their correlations were obtained with SPSS 18. The models adjusted in MLWin are as follows:

Level 1:

$$y_{ij} = \beta_{0ijk} Cons + \beta_1 GENERO_{ijk} + \beta_2 LENGUA_{ijk} + \beta_3 ZNSC1_{ijk} + \beta_4 ZNE_{ijk} + \beta_5 NSE\_ESCUELA_{ji} + \varepsilon_{ij}$$

Level 2:

$$\beta_{0jk} = \beta_{00k} + \mu_{ojk}\,;\ \beta_{1jk} = \beta_{10k} + \mu_{1jk}\,;\ \beta_{2jk} = \beta_{20k} + \mu_{2jk}\,;\ \beta_{3jk} = \beta_{30k} + \mu_{3jk}\,;$$

$$\beta_{4jk} = \beta_{40k} + \mu_{4jk}\,;\ \beta_{5jk} = \beta_{50k} + \mu_{5jk}$$

Level 3:

$$\beta_{00k} = \gamma_{000} + v_{ook}\,;\ \beta_{10k} = \gamma_{100} + v_{10k}\,;\ \beta_{20k} = \gamma_{200} + v_{20k}\,;\ \beta_{30k} = \gamma_{300} + v_{30k}\,;$$

$$\beta_{40k} = \gamma_{400} + v_{40k}\,;\ \beta_{50k} = \gamma_{500} + v_{50k}$$

Where:

y_{ijk} is the result variable and i = 1,..., ni is every student, j = 1,..., nj every of the school, k= 1...., nk every of the state; e_{ijk}, μ_{ojk}, V_{0k} are the errors of every level.

β_{njk} = it is the variables included in the models.

β_{0jk} = it is the average score for the student 0 of the school j and state k, it is compound by β_{00k} it is the average of the intercept of the school, and μ_{0jk} the random error.

β_{1jk} = it is the average of the slope, it shows the variability between school j and state k, it is compound by β_{10k} is equal the average of the slope by the schools, and μ_{1jk} the random error.

β_{00k} = it is the average score by the intercept for the three level, it is compound by γ_{000} is the general average and v_{ook} random error.

β_{10k} = it is the average slope of the three levels and it is compound by γ_{100} the average of intercept and v_{1ok} the random error.

For the model introduces the variables in the fixed part after that, there is variation in the average and slopes.

The stability was calculated for cohorts with longitudinal multilevel model, we used MLwiN 2.3. The model is the following:

$$y_{ij} = \beta_{0ij} Cons + \beta_{1ij} pension_{1j}$$

$$\beta_{0ij} = \beta_0 + \mu_{oj} + e_{oij}$$

$$\beta_{1ij} = \beta_1 + \mu_{1j} + e_{1ij}$$

$$\begin{bmatrix} \mu_{oj} \\ \mu_{1j} \end{bmatrix} \text{N}(0, \Omega u): \Omega u = \begin{bmatrix} \sigma^2_{u0} & \\ \sigma^2_{u01} & \sigma^2_{u1} \end{bmatrix}$$

$$\begin{bmatrix} e_{oij} \\ e_{1ij} \end{bmatrix} \text{N}(0, \Omega e): \Omega e = \begin{bmatrix} \sigma^2_{e0} & \\ \sigma^2_{e01} & \sigma^2_{e1} \end{bmatrix}$$

Where:

Y_{ij} = it is the score of the test by the time i of the school j.

β_{0j} = it is the average score according to the time for the school *j*.

β_{1j} = it is the slope of the variable tuition.

σ^2_{e1} = it is the variability of first models, time.

σ^2_{u1} = it is the variability of second level by school.

4. Results

The results of consistency are shown in Table 2. The values for consistency are high for every subject except from 2000 and 2001. In 2000, the correlations are high for the association among language, social and natural science. The correlation is low between mathematics and language but moderate for the other subjects. For 2003, the correlation is high among language, social and natural science and moderate for the other subjects.

https://doi.org/10.15405/ejms(2421-8251).2016.1.2
eISSN: 2421-8251 / Corresponding Author: Olga Rosalba Rodríguez-Jiménez
Selection & Peer-review under responsibility of the Editors

Table 2. Consistency among subjects

	Years						
Areas	2000	2001	2002	2003	2008	2009	2010
Mathematics-language	0,349	0,4812	0,5763	0,4321	0,7063	0,7315	0,7536
Mathematics - natural science	0,4921	0,4898	0,6645	0,5008	0,7414	0,8031	0,7579
Mathematics -Social science	0,4664	0,5123	0,6725	0,4595	0,7101	0,7439	0,786
Language-naturalscience	0,7261	0,7363	0,7521	0,8026	0,7588	0,7781	0,831
Language-Social science	0,7789	0,8026	0,8736	0,8281	0,7792	0,7862	0,7649
Natural-Social sciences	0,7822	0,7745	0,8204	0,8155	0,7912	0,822	0,8052

On the other hand, the results of stability showed that there is meaning difference for the time and the tuition. In table 3, we noticed that the average for mathematics is 42,73 with statistics difference among schools by the time and tuition (μ_{1j} =0,409 y ee=0,043). For language the average score is 45,645 being the highest, there is difference among schools by the time and tuition (μ_{1j} =0,092 y ee=0,039).

In the case of natural science and social science the average score is 44,052 and 42,34 respectively with difference by time and tution like the other two subjects (μ_{1j} =0,126 and ee=0,030; μ_{1j} =0,075 y ee=0,029).

In summary, the variability showed by the score is not the same for the subjects by the time, this has impact on the school effect.

Table 3. Stability of four subjects for seven years

	Mathematics		Language		Natural science		Social science	
	B	Ee	B	Ee	B	ee	B	Ee
Fixed part								
Intercept	42,730	0,030	45,645	0,039	44,052	0,027	42,341	0,031
Tuition	0,691	0,020	0,960	0,017	0,620	0,016	0,713	0,015
Random part								
Schools	2,992	0,085	5,837	0,148	2,711	0,069	3,493	0,091
Time	5,394	0,042	3,922	0,031	1,405	0,011	3,002	0,023
Difference of probability*	33163,684		23978,059		21982,061		23550,296	

Note: *Probability maximun with P< 0,01.

5. Discussion

5.1. Consistency

The results show that the school effectiveness is equal for all the subjects and have implications for the organization. In high school, the consistency is moderate similar to Cervini's (2010) research in Argentina and Zorrilla's (2008) in Mexico. However, this study's findings contradict the results obtained by Cuttance (1987), Thomas and Mortimore (1996) and Thomas and Golsdtein (1994) who reported moderate correlations among subjects and with the results of Golsdtein and Rasbash (1993) who showed low correlations close to 0.29. These results allow for an assumption that consistency is similar to the other

Latin American countries. Hence, it is possible to affirm that these results are due to the similar characteristics of the school. It is important to mention that only mathematics posted a significant change in school effectiveness.

The consistency in this research is high with values more than 0.7, except for mathematics which is moderate and low for 2000 in particular, with language. These results are similar to those obtained by Bosker and Scheerens (1989), Cervini (2010), Zorrilla (2008) and Reynolds et al. (2011).

In summary, the Colombian school effect, either positive or negative, is similar across the board for every subject as consistency was established among subjects except for matemathics because the schools that have good results in this subject are not good at the other subjects (Murillo, 2007a).

5.2. Stability

The results for stability showed that there is change due to time for every subject. Consequently, it is possible to assume that the Colombian school is not stable, as the school effect is not stable over time. In addition, these results depend on the socio-economic status of the school. These findings are similar to Mangan, Pugh, and Gary (2005) who affirmed that there are school changes which lead to instability over time. In the current research, it was found that along the ten-year period, ascendant and descendant tendencies were noticeable. This behavior is evidence of instability.

With regard to the results for stability, it is difficult to compare the results of this study with those of emblematic studies such as Thomas et al. (1997) who showed that stability depended on the subject or the study of Thomas, Peng, and Gray (2007) who found that the 55% of schools are effective over time, with many unexpected changes. For this reason, only in a few schools, is it possible to find improvement.

The results for stability and consistency\show that although there is similarity in the teaching-learning process among subjects, this changes from one year to the next. The variability could be the answer to the arrangement of the external circumstances (Goldstein, 1997). Despite changes in institutional or governmental policies, they influence all the subjects in the same way. These results show the need to consider the dynamic models (Kyriakides & Creemers, 2010). Like other studies in Latin America, the results of this research show:

> … the important repercussion in the index of effectiveness of the schools... if the school effect is not consistent, we would analyze the effectiveness in a different measure of the results; if they are not stable, it will be necessary to have the index measure for some years consecutive… (Murillo, 2007a, p. 83)

5.3. Quality and equity education and social equity

The research about school effect informs us about the effectiveness of school. This shows the results of educational processes that are ahead in the schools and reflect the politics of state or country. These results are the principal tools for informing about the quality of education.

The quality of education comprises five dimensions: relevance, pertinence, equity, effectiveness and efficiency (Blanco, 2008). Relevance and pertinence refer to the aim and sense of education. The other dimensions have to do with guaranteeing access and permanence inside the educational system like equity and the achievements of educational goals. They also are related to the way money is invested for education. The studies about consistency and stability are useful for providing information about these dimensions.

The results of this research show that in Colombia, schools are not very successful in producing good results in different subjects and guaranteeing their stability. This will impact on the quality of education. If Colombia wants to become developed in the way that the UNESCO (2012) suggests, the country should work towards decreasing the differences in the quality of the education because these have a considerable effect on opportunities for employment and on national development. In conclusion, the country must be seriously concerned about putting in place fairer, participatory and equitable societies that allow for a harmonious and prosperous coexistence among citizens.

Acknowledgements

The author(s) declare that there is no conflict of interest.

References

Becerra-Avella, M. O. (2010). *Comparación del análisis factorial múltiple (AFM) y del análisis en componentes principales para datos cualitativos (Prinqual), en la construcción de índices* [Comparison of multiple factor analysis (AFM) and principal components analysis for qualitative data (Prinqual), in the construction of indices]. (Master's Thesis). Universidad Nacional de Colombia, Bogotá Colombia.

Blanco, R. (2008). Eficacia escolar desde el enfoque de la calidad de la educación [School effectiveness from the quality of education approach]. *En OREALC/UNESCO, Eficacia escolar y factores asociados en América Latina y el Caribe* (pp. 7-16). Santiago: OREALC/UNESCO.

Bosker, R. J., & Scheerens, J. (1989). Issues in the interpretation of the results of schools effectiveness research. *International Journal of Educational Research, 13*(7), 741-751. https://doi.org/10.1016/0883-0355(89)90025-6

Cervini, R. (2010). El 'efecto escuela' en la educación primaria y secundaria: el caso de Argentina [The 'school effect' in primary and secondary education: the case of Argentina]. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, 8,* 8-25.

Cuttance, P. F. (1987). *Modelling variation in the net effects of schools on pupil outcomes.* Edinburgh: CES.

Fitz-Gibbon, C., Tymms, P., & Hazlewood, A. (1990). Performance indicators and information system. In D. Reynolds, B. P. M. Creemers, & D. Peters (Eds.), *School effectiveness and improvement*. Groningen: GION.

Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement, 8*(4), 369-395. https://doi.org/10.1080/0924345970080401

Goldstein, H., & Rasbash, J. (1993). A multilevel analysis of school examination results. *Oxford Review Of Education, 19*(4), 425-440. https://doi.org/10.1080/0305498930190401

ICFES. (2005). *Nueva estructura de los exámenes de estado para ingreso a la educación superior y validación del bachillerato académico en un solo examen* [New structure of the state exams for admission to higher education and validation of the academic baccalaureate in a single exam]. Bogotá: ICFES.

ICFES. (2011c). *Examen de Estado de la educación media Resultados del periodo 2004 – 2009.* Bogotá: ICFES.

Kyriakides, L., & Tsangaridou, P. (2008). Towards the development of generic and differentiated models of educational effectiveness: a study on school and teacher effectiveness in physical education. *British Educational Research Journal, 34*(6), 807-838. https://doi.org/10.1080/01411920802041467

Kyriakides, L. (2004). Differential School Effectiveness in Relation to Sex and Social Class: Some Implications for Policy Evaluation. *Educational Research and Evaluation: An International Journal on Theory and Practice, 10*(2), 141-161. https://doi.org/10.1076/edre.10.2.141.27907

Kyriakides, L., & Creemers, B. (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. https://doi.org/10.1080/03054980701782064

Kyriakides, L., & Creemers, B. (2010). Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness and School Improvement, 21*(4), 409–427. https://doi.org/10.1080/09243453.2010.512795

Luyten, H. (2004). School effects: stability and malleability. Enschede: University of Twente.

Mandeville, G. K. (1988). School Effectiveness Indices Revisited: Cross-Year Stability. *Journal of Educational Measurement, 25*(4), 349-356. https://doi.org/10.1111/j.1745-3984.1988.tb00313.x

Mandeville, G. K., & Anderson, L. W. (1987). The Stability of School Effectiveness Indices across Grade Levels and Subject Areas. *Journal of Educational Measurement, 24*(3), 203-216. https://doi.org/10.1111/j.1745-3984.1987.tb00275.x

Mangan, J., Pugh, G., & Gray, J. (2005). Changes in examination performance in English secondary schools over the course of a decade: searching for patterns and trends over time. *School Effectiveness and School Improvement, 16*, 19-50. https://doi.org/10.1080/09243450500114843

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters: The junior years*. Somerset: Open Books.

Murillo, F. J. (2005a). *La investigación sobre eficacia escolar* [Research on school effectiveness]. Barcelona: Octaedro.

Murillo, F. J. (2005b). ¿Importa la escuela? Una estimación de los efectos escolares en España [Does school matter? An estimate of school effects in Spain]. *Tendencias Pedagógicas, 10*, 29-45.

Murillo, F. J. (2007a). *Investigación Iberoamericana sobre Eficacia Escolar* [Ibero-American Research on School Efficiency]. Bogotá: Convenio Andrés Bello.

Murillo, F. J. (2007b). School effectiveness research in Latin America. In T. Townsend (Ed.), *International Handbook of School Effectiveness and Improvement* (pp. 75-92). New York, NY: Springer. https://doi.org/10.1007/978-1-4020-5747-2_5

Murillo, F. J. (2008). Enfoque, situación y desafíos de la investigación sobre eficacia escolar en América Latina y el Caribe [Approach, situation and challenges of research on school effectiveness in Latin America and the Caribbean]. *Eficacia escolar y factores asociados en América Latina y el Caribe* (pp. 17-48.). Santiago de Chile: OREALC/UNESCO.

Murillo, F. J., & Hernández-Castilla, R. (2011). Efectos escolares de factores socio-afectivos. Un estudio Multinivel para Iberoamérica [School Effects of Socio-Affective Factors. A Multilevel Study for Latin America]. *Revista de Investigación Educativa, 29*(2), 407-427.

Opdenakker, M.-C., & Van Damme, J. (2000). Effects of schools, teaching staff and classes on achievement and well-being in secondary education: Similarities and differences between school outcomes. *School Effectiveness and School Improvement, 11*, 165-196. https://doi.org/10.1076/0924-3453(200006)11:2;1-Q;FT165

OREALC/UNESCO. (2010). *Factores asociados al logro cognitivo de los estudiantes de América Latina y el Caribe* [Factors associated with the cognitive achievement of Latin American and Caribbean students]. Santiago de Chile: OREALCUNESCO.

Postlethwaite, T. N. (1994). Monitoring and Evaluation in Different Education Systems. In A. C. Tuijnmanand, & T. N. Postlethwaite (Eds.), *Monitoring the Standards of Education* (pp. 23-46). Oxford: Pergamon.

Reynolds, D., Sammons, P., De Fraine, B., Towsend, T., & Van Damme, J. (2011). Educational Effectiveness Research (EER): A State of the Art Review. Comunicación presentada en la reunión anuel del *International Congress for School Effectiveness and Improvement*, Cyprus.

Rodríguez-Jiménez, O. R., & Murillo, F. J. (2011). Estimación del efecto escuela para Colombia [Estimation of the school effect in Colombia]. *Magis, Revista Internacional de Investigación en Educación, 3*(6), 299-316.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: secondary schools and their effects on children*. London: Open Books.

Sammons, P., Nuttall, D., & Cuttance, P. (1993). Differential school effectiveness: results from a reanalysis of the Inner London Education Authority's Junior School Project Data. *British Educational Research Journal, 4*, 381-405. https://doi.org/10.1080/0141192930190407

Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. London: Elsevier.

Thomas, S. (2001). Dimensions of secondary school effectiveness: Comparative analyses across regions. *School Effectiveness and School Improvement, 12*, 285-322. https://doi.org/10.1076/sesi.12.3.285.3448

Thomas, P., & Golsdtein, H. (1994). *The Analysis of 1992 GCSE Results*. London: Association of Metropolitan Authorities.

Thomas, S., & Mortimore, P. (1996). Comparison of value-added models for secondary school effectiveness. *Research papers in Education, 11*(1), 5-33. https://doi.org/10.1080/0267152960110103

Thomas S., Peng, W. J., & Gray, J. (2007). Value added trends in English secondary school performance over ten years. *Oxford Review of Education, 33*, 261-298. https://doi.org/10.1080/03054980701366116

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. https://doi.org/10.1080/0924345970080201

UNESCO. (2012). *Informe de seguimiento de la EPT en el mundo: los jóvenes y las competencias* [EFA Global Monitoring Report: Youth and Skills.]. París: UNESCO.

Willms, J. D., & Raudenbush, S. W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement, 26*, 209-232. https://doi.org/10.1111/j.1745-3984.1989.tb00329.x

Zorrilla, M. (2008). *La escuela secundarios mexicana. Un estudio multinivel de los efectos escolares y sus propiedades científicas* [The Mexican high school. A multilevel study of school effects and their scientific properties]. Doctoral Dissertation. Universidad Anahuac. México.