

ARTICLE

ASSOCIATION OF TEACHING PRACTICES AND STUDENT PERFORMANCE IN LARGE SCALE EVALUATION OF THE STATE OF ESPÍRITO SANTO

NAIRA MUylaERT^{1*}

ORCID: <https://orcid.org/0000-0001-5161-0501>

ANA LUISA MARLIÈRE CASELA^{II**}

ORCID: <https://orcid.org/0000-0002-3321-0325>

LUIZ VICENTE FONSECA RIBEIRO^{III***}

ORCID: <https://orcid.org/0000-0002-0000-561>

WAGNER REZENDE^{IV****}

ORCID: <https://orcid.org/0000-0003-1617-282X>

ABSTRACT: The present study is in line with the field of investigation of school effectiveness, because the main objective is to investigate the association between student performance in a large-scale evaluation and the scale of Attitudes and Teaching Practices (IPP) - defined as one of the factors of school effectiveness and that refers more directly to the teacher's performance. Through a quantitative approach, the study made use of a statistical model called HLM to observe the association between IPP and the performance in Portuguese Language and Mathematics of the 9th grade students, evaluated by PAEBES - a large-scale evaluation of the state of Espírito Santo. The main results corroborate other research findings and relate to the negative effect of the reprobation and the positive effect of IPP on the performance of both disciplines.

Keywords: Teaching Practices. School Effectiveness. School performance. HLM.

¹ Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.

^{II} Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brasil.

^{*} PhD in Human Sciences: Education. Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Postdoctoral student in Educational policies of the Education Department of PUC-Rio. Member of the Laboratory of Education Evaluation (LAEEd / PUC-Rio). E-mail: <naira.muylaert@hotmail.com> .

^{**} Master in Psychology. Universidade Federal de Juiz de Fora (UFJF). Analyst of Education Evaluation at the Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF). E-mail: <analuisacasela@caed.ufjf.br> .

^{***} PhD in Social Sciences, form Universidade Federal de Juiz de Fora (UFJF). Supervisor of Research Support in Educational Evaluation at the Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF). E-mail: <Iribeiro@caed.ufjf.br> .

^{****} PhD in Social Sciences, Universidade Federal de Juiz de Fora (UFJF). Results Delivery and Professional Development Supervisor at the Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF). E-mail: <wagner@caed.ufjf.br> .

ASSOCIAÇÃO DO ÍNDICE DE ATITUDES E PRÁTICAS PEDAGÓGICAS AO DESEMPENHO DOS ESTUDANTES NA AVALIAÇÃO EM LARGA ESCALA DO ESTADO DO ESPÍRITO SANTO

RESUMO: O presente estudo alinha-se ao campo de investigação da eficácia escolar na medida em que tem como objetivo principal investigar a associação entre o desempenho dos estudantes em uma avaliação em larga escala e o índice Atitudes e Práticas Pedagógicas (IPP) – definido como um dos fatores de eficácia escolar e que se refere, de forma mais direta, à atuação do professor. Por meio de uma abordagem quantitativa, o estudo fez uso de uma modelagem estatística denominada HLM para observar a associação entre o IPP e o desempenho médio em Língua Portuguesa e Matemática dos estudantes do 9º ano do Ensino Fundamental avaliados pelo PAEBES – avaliação em larga escala do estado do Espírito Santo. Os principais resultados corroboram outros achados de pesquisa e dizem respeito ao efeito negativo da reprovação e do efeito positivo do IPP sobre o desempenho de ambas as disciplinas.

Palavras-chave: Práticas Pedagógicas. Eficácia Escolar. Desempenho Escolar. HLM.

INTRODUCTION

Throughout the school education process, the role of the teacher plays a central role, as he/she interacts directly with the students, coordinating the daily educational activities and mediating the teaching and learning processes. Their performance has a great discretionary power in relation to the execution of educational policies, as they decide how to manage the class, what methods they will use to present ideas and develop skills and competences, ways to communicate with students and other decisions that they make on a daily basis that directly and indirectly affect the quality of education (HANUSHEK, 1986). It is mainly in the classroom and in the teacher-student interaction that learning happens. It is not surprising, therefore, the great attention that public administrators and researchers in education dedicate to the performance of teachers and the pedagogical practices conducted by them.

With the expansion of large-scale evaluations and the results produced from them, the debate about the school effect and the teacher effect – that is, the capacity that school actions (school effect)

and that the action of teachers (teacher effect) has to produce results in terms of learning, or performance – has expanded. These school actions and, more specifically, the actions performed by teachers are defined by the field of study of school effectiveness as factors that may or may not be associated with performance. By knowing these factors and verifying their effect on performance, it is possible to propose policies that aim to improve the effectiveness of school actions and, especially, the performance of teachers, which can generate positive impacts on the performance results of large-scale evaluations (MORICONI, 2012).

In general, the national and international literature on teacher effectiveness addresses aspects related to the characteristics of teachers, but which are external to the classroom, such as experience in teaching, level of education and performance in the selective processes of career entry. The interest of the studies in these aspects is due to the wide availability of data and also the importance they have in the decision making regarding the policies of the teaching profession. Although some studies (RIVKIN, HANUSHEK E KAIN, 2005; CLOTFELTER, LADD E VIGDOR, 2007; ROCKOFF, 2004; DARLING-HAMMOND, 2005; BARROS ET AL, 2001; ALBERNAZ, FERREIRA E FRANCO, 2002, SOARES, 2003) point out positive relationships between these characteristics and student performance, the associations of these factors to learning are small (HANUSHEK, 1986).

While some researchs find weak or no association between the observable characteristics (schooling, salary, experience) of teachers and performance, other studies focus on the effectiveness of teacher performance - a set of unobservable characteristics, called latent trait.

In the American context, the literature review by Nye, Konstantopoulos and Hedges (2004) presents a series of studies based on different student samples, which indicate that 7% to 21% of the variation of the grades in the standardized tests are associated to the variation in the effectiveness of teachers. Another compilation by Hanushek and Rivkin (2004) analyzes eleven studies with data of a longitudinal nature and indicate variations from 0.08 to 0.36 standard deviation as a function of teacher effectiveness in the student's results. According to Wright, Sanders and Horn (1997), other studies using longitudinal data from the Tennessee Value-Added Assessment System (TVASS) show that teacher effectiveness is the major determinant of student academic achievement, including the effects of socioeconomic status.

These studies are based on the investigation of the effects of teachers looking at the differences in performance growth in standardized tests and depart from the definition that a good teacher is one who can achieve higher growth in terms of his/her students' grades, while a bad teacher is the one that produces smaller growths of learning (MORICONI, 2012).

A study by Sammons (1999) identifies some key characteristics of effective schools and is associated with good school performance. Some of these characteristics relate directly to the teacher. One of them is defined by Sammons (1999) as a *consistent practice* and refers to the consistency with which teachers develop their practices and follow common, collectively defined and informed approaches to all school actors. In other words, this means that when there is unity of purpose, the performance of teachers becomes more consistent and this can positively contribute to student performance.

Another aspect concerns the clarity of teaching objectives, which includes elements such as effective organization (prepare/plan class assignments in advance), clarity of objectives (when teachers clearly explain the objectives of the lesson and refer to them in the duration of the lesson), and adaptive practice (perception that students are different, have different rhythms and forms of learning and with this, the teacher makes use of appropriate teaching strategies. It requires flexibility to modify and adapt their class styles).

Sammons (1999) also stresses the importance of high expectations for learning, of communicating these expectations to students, and of providing positive incentives in either discipline patterns (developing feelings of belonging and participation) or feedback (success recognition). Still with regard to expectations, Cardelli and Elliot (2012), analyzed a success story of a school in risk area in Rio de Janeiro. They pointed out the expectations of teachers in relation to the performance of students as one of the factors responsible for the good results at school.

Finally, well structured classes are also characterized as an important element associated with good performance. Drawing on several research findings, Sammons (1999) defines well structured classes as practices that can keep students' attention. According to the literature, some of these practices are: a) techniques for asking open, pertinent and challenging questions; b) make clear what has to be learned; c) offer diversified materials; d) pass and correct homework; e) teach for all students. It is clear that there are other features that permeate the teaching-learning process and promote satisfactory

learning. These are just some practices, pointed out in the literature, that are associated with performance.

In Brazil, studies on school effect and teacher effect (MACHADO SOARES, 2003; SOARES, 2004; SOARES e ALVES, 2007; OLIVEIRA, 2012; MORICONI, 2012) are historically more recent and were carried out under the influence of foreign research, especially those carried out in the US context. In addition, the questionnaires of large-scale evaluations (national and state) applied in the Brazilian territory were constructed based on foreign literature. That is, it is from this theoretical foundation that the constructs, or latent traits measured by the contextual questionnaires of the evaluations were defined and hence the importance of reference to them.

The main characteristics of the teacher's work mentioned above served as reference for the construction of the items related to the Pedagogical Attitudes and Practices (IPP) block of the questionnaire applied to students and teachers participating in the large-scale educational evaluation of the State of Espírito Santo. From these items it was possible to construct the IPP indicator that is configured as object of the present work.

Thus, in line with the field of research on school effectiveness, this study aims, through a quantitative approach, to investigate the association between performance in Portuguese and Mathematics of the students participating in the large-scale evaluation of the State of Espírito Santo and the Pedagogical Attitudes and Practices (IPP) construct—defined as a factor that is positively associated to performance and which refers more directly to the teacher's performance.

The work is divided into 3 sections, in addition to this introduction. In the next section we present the analysis strategies of the data and the procedures used to carry out this study. The third section presents and discusses the main results found and, finally, the final considerations.

ANALYSIS STRATEGY

In order to empirically verify the association between the effective performance of teachers and students, we use the contextual and performance data of the Basic Education Evaluation Program of Espírito Santo for 2015 (PAEBES/2015). To do so, we use the measures of proficiency¹ in Portuguese Language and in Mathematics of the students of the 9th year of elementary school and the Indicator of Pedagogical Attitudes and Practices (IPP). To

guarantee the reliability and validity of the results, we used some measures of socioeconomic control in the analysis, namely: gender, color/race, reprobation and socioeconomic level of students.

We know from numerous researches (SOARES e ALVES, 2003; CRAHAY, 2013) that the students' gender, race and school trajectory make a difference in their performance, as well as in the teaching and learning process, being one of the main aspects emphasized in the studies on the relations between social inequalities and educational inequalities. The literature demonstrates how gender, race and trajectory, characteristics often considered "natural" to individuals, become socially relevant attributes for performance in standardized tests. This information is raised in this study, therefore, in order to see the real association of pedagogical practices and student performance. The individual characteristics are raised through items in the students' contextual questionnaire. A factor related to the school trajectory refers to the reprobation. Strictly speaking, repetition and reprobation are different concepts. In effect, repetition refers to the school practice of retaining students in a certain school year. According to Correa (2013, p.20), the repeater "would be any student who enrolled in the same grade as in the previous year, regardless of the reason that led him to do it." The repetition refers, therefore, to students who attend the same school year twice, for whatever reason. So, for example, the student who left school and the following year returned to the same school stage he did the year before is considered a repeater, but not a disapproved. On the other hand, reprobation is the school practice of retaining the student for unsatisfactory learning. That is, the student does not advance in his school career because he did not learn the contents of the academic stage he attended, leading him to do the same step again. The information about reprobation, in this study was raised by an item in the students' contextual questionnaire.

On the other hand, the socioeconomic status is a latent construct, that is, it can not be directly observed by us as the individual characteristics. For this reason, we need a strategy to get information about observable characteristics that reveal something about this dimension. In the context of educational evaluation, what we can do is to include questions in the questionnaires answered by the students, about the goods, services and characteristics of the houses. These questions include economic aspects (consumer, access to public services, etc.), and social and cultural aspects (access to culture, parental schooling, etc.), which allowed the creation of the

Socioeconomic Index (ISE) – a measure used to numerically represent the situation of students and their families. The ISE was calculated based on the Item Response Theory (TRI), items calibration and ISE calculation were produced using the software Parscale (1996), following the model for graduated responses of Samejima (1969). This model assumes that the response options of the items are ordered and cumulative, that is, the classification at a certain level implies achieving the necessary to the previous level and comply with something else. For example, if an item where the respondent answers about the number of cars his family owns the answer is 2, that means that the previous level (one car) has been met and there is still a gain from that previous level (another automobile). This differentiation between levels can have a character quantitatively and qualitatively. From the individual's answers to the items, the parameters were calibrated using the Maximum Likelihood (ML) method. These parameters allow the model to define the probabilities of respondents for each category in each item. Taking the default response and the set of probabilities for all items, the index for each questionnaire was calculated using the Mean to Posteriori (EAP) method.

The performance of teachers in the classroom has multiple dimensions that can not be fully and holistically captured in quantitative approach studies. This is mainly due to the technical limitation of the data collection instrument itself - the questionnaires. These instruments present items² that seek to capture information about the perception of the respondents (students and teachers) about some dimensions of the construct, defined here as Pedagogical Attitudes and Practices (IPP). Based on the answers to the questionnaire items, it is possible to construct an indicator that synthesizes the construct, or latent trait.

The items referring to the IPP construct were answered by the students (5th and 9th years of elementary school and 3rd year of high school) and by the teachers. They are *Likert type* items (DALMORO, M.; VIEIRA, K., 2013), with options for gradual answers: I totally agree, I agree more than I disagree, I disagree more than I agree and I totally disagree. The items were staggered non-parametrically (FIELD, A., 2009; SIEGEL, S. AND CASTELLAN Jr. J.N., 2006), with values from 1 to 4, depending on the direction of the item in relation to the construct. Then, also non-parametrically, we used the anchorage vignette methodology (KING *et al*, 2004; KING AND WAND, 2007; HOPKINS AND KING, 2010; KYLLONEN AND BERTLING, 2013) to restagger the items.

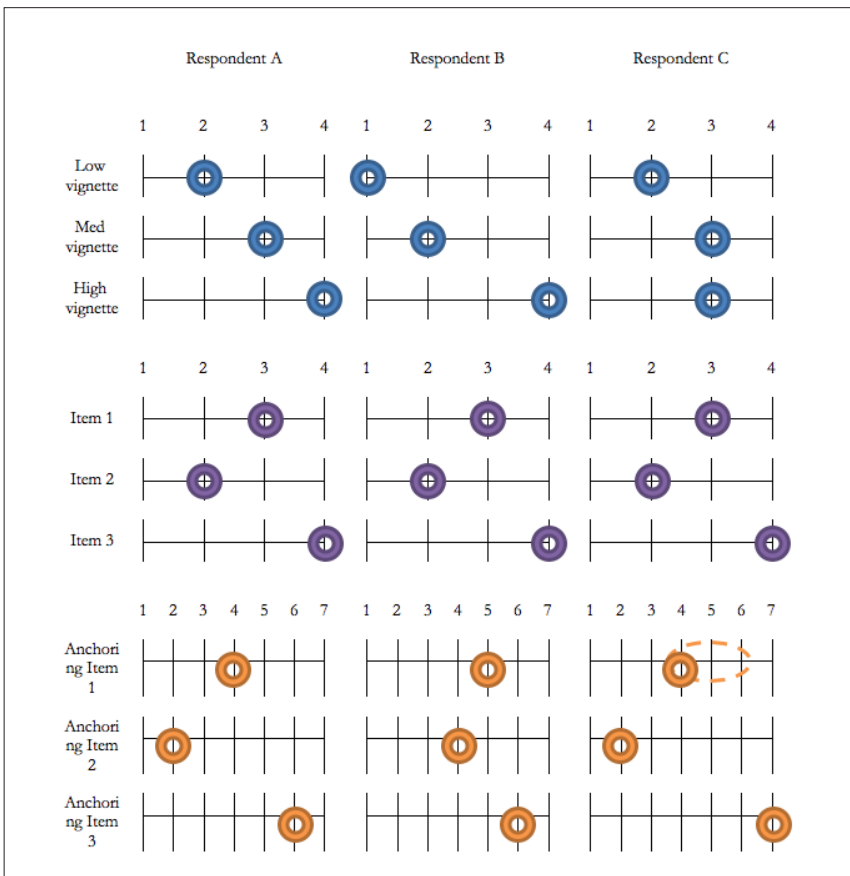
According to this methodology (KING *et al.*, 2004; KING AND WAND, 2007), to improve the frame of reference for respondents and reduce differential item functioning (“DIF”), researchers should construct items that concretely describe the assessed dimension, so respondents can “position” themselves on the continuum represented by the vignettes. Differential item functioning, or “DIF”, is the most modern terminology for interpersonal incomparability of responses, considering the variation in the understanding of the meaning of the items (KING *et al.*, 2004). According to King *et al.* (2004):

Our key idea, in addition to following the venerable tradition of trying to write clearer questions that are more comparable, is a method of directly measuring the incomparability of responses to survey questions, and then correcting for it. We ask respondents for self-assessments of the concept being measured along with assessments, on the same scale, of each of several hypothetical individuals described by short vignettes. We create interpersonally comparable measurements by using answers to the vignette assessments, which have actual (but not reported) levels of the variables that are the same for every respondent, to adjust the self-assessments. Our adjustments can be made with simple calculations (straightforward recode statements) or with a more sophisticated statistical model that has the advantage of lowering data collection costs. (KING *et al.*, 2004, p. 192)

The vignettes, therefore, function as anchor points on an ordered scale, where items are assigned a value of 1 if the item is classified below the vignette with the lowest profile presented, with a value of 2 if the item was rated equal to this, 3 if it is between this and the median profile, 4 if it is equal to the median, 5 if it is between the median and the high profile, 6 if it is equal to the high profile and 7 if it is classified above the high profile. This recoding, based on vignettes, results in a “DIF-free” response, with easily interpretable ordinal levels. There are two cases of special interest when the self-reported evaluation of vignettes does not follow the expected order: “ties” and “violations”. Ties occurs when the respondent chooses the same response category for the vignettes, and violations occurs when the respondent chooses a higher response category for the vignette(s) that represents a lower profile. Following one of the method’s for scoring PISA 2012 Likert items described in OECD (2013), all violations were considered as “broad ties” and among the values derived from the recoding given by this broader spectrum, the choice was made at the lowest value, or be it at the lower limit of the range of possibilities. The choice for the lower limit score created the greatest variation and allowed the best differentiation between the individuals. For example, let’s say that an respondent rated the vignettes in the following order: low, high, med.

This classification was considered as follows: low, {med, high}, with the brackets indicating that the same classification is assigned to the med and high vignettes. A graphic illustration of the vignette-based recoding procedure for three examples with and without ties is given by Figure 1. The three hypothetical students in this example provided exactly the same answers to the three items reported but differ in their responses to the vignettes. As a result, scores on anchored items also differ. More details on the scoring approach are provided in King *et al* (2004), King and Wand (2007) and OCDE (2013).

FIGURE 1. Representation of the anchoring of items for three hypothetical respondents.



Source: Elaboration by authors based on OECD, 2013.

Below are the vignettes and items present in the questionnaires applied to students and teachers participating in the PAEBES evaluation in 2015.

Vignettes presented to students

- Alberto's teacher uses different resources to teach: board, TV, music, etc. She also uses the textbook and appears with different activities for classroom work. (High vignette).
- Cassia's teacher uses more of the textbook, sometimes he passes some video or uses the computer lab, and passes activities to the class in the classroom. (Med vignette).
- Olavo's teacher works with the textbook in the classroom, from time to time she passes activities to her students, but the exercises are very similar. (Low vignette).

Items presented to students of the 5th ES

- My teachers are interested in student learning.
- My teachers use different ways to present the subject (video, music, computer programs, etc.).
- My teachers respond to questions that come up as they teach.
- My teachers correct homework in class with the students.

Items presented to students of the 9th ES and 3rd HS

- My teachers use different ways to present the subject (video, music, computer programs, etc.).
- My teachers correct the exercises in the classroom with the students.
- My teachers respond to questions that come up as they teach.
- My teachers correct homework in class with the students.
- My teachers return the research papers if I do not do it right.
- My teachers review the subject for the exams.
- My teachers read the tests with the students and answer the questions (about the test) before we begin to do it.
- My teachers have good teaching practices.

Vignettes presented to teachers

- Pedro's teacher uses different resources to teach (board, TV, music, etc.), also uses the textbook, and proposes different activities for classroom work. (High vignette).
- Paula's teacher uses the textbook more. Sometimes she passes some video or uses the computer lab and passes classroom activities to the class. (Med vignette).
- Olavo's teacher works with the textbook in the classroom, from time to time he passes activities to his students, but the exercises are very similar. (Low vignette).

Items presented to teachers

- I use different ways to present the subject (video, music, computer, etc.).
- I use the textbook very much in class.
- I correct the exercises in the classroom with the students.
- I respond to questions that come up as I teach.
- I correct homework in class with the students.
- I return research papers if students do not do it right.
- I make it clear to my students that learning is very important and I always try to motivate them.
- I talk to the students about the subject of the tests before applying them.
- I read the tests with the students and I answer questions about the test before they begin to do it.
- I use the results of the evaluations I develop to review my lesson plan.
- I give attention to what is programmed, independent of the class.
- I am interested in student learning.

In order to verify if the questions that deal with pedagogical attitudes and practices are a measure that synthesizes the construct, we evaluate the items through their correlations, the *Cronbach's Alpha* and the main components analysis. These analyzes were conducted separately using the microdata bases of the students at each stage and the teachers. We adopted, as recommended by the literature (FIELD, 2009), minimum parameters: the correlations between items higher than 0.3, *Cronbach's Alpha* greater than 0.5, the *Kaiser-Meyer-Olkin* (KMO) suitability measure greater than 0.8, the percentage of the variance explained by the first factor greater than 40%, the eigenvalue measure greater than 1, and the slope diagram showing the preponderance of a single factor. In all cases, these criteria were satisfied and, therefore, we extracted the first factor to represent the perception of the actors. Subsequently, the scores were aggregated by the mean in the respective school. To aggregate teacher and student scores of the different stages, we used the weighted average number of students evaluated at the school.

After the construction of the IPP, we used hierarchical linear modeling to estimate the influence of this index on student proficiency. Hierarchical models are considered the most appropriate in the literature on school effectiveness, since they are able to distinguish,

within the same school and between schools, the difference in student performance. Thus, it is possible to use individual characteristics of the students (i.e., factors external to the school, such as socioeconomic status, color/race, and gender) to measure how much they affect their performance within the school and also to use characteristics of the schools (i.e., factors internal to the school, such as pedagogical attitudes and practices) to measure more precisely how much school factors impact the performance of these students.

In addition to the individual factors of the students' level, such as gender, color/race, socioeconomic level and reprobation, the model also allows to evaluate at school level how much their social composition influences the expected proficiency: percentage of girls, percentage of whites, percentage of reprobated and average socioeconomic level of the school. And, as the main objective, the model also allows us to verify to what extent the perception that the school actors (students and teachers) have about pedagogical attitudes and practices influences the explanation of average proficiency. From the presentation of the results it is possible to say which factors have an effect on the performance, as well as the direction (positive or negative) and the force of that influence (high or low).

RESULTS & DISCUSSION

The variables that composed the model can be divided into three groups according to their coding form. The first refers to proficiency information in Portuguese Language and Mathematics, characterized, from the statistical point of view, as continuous variables on a natural scale. In PAEBES, these scales (one of Portuguese Language and one of Mathematics) have an average of 250 and a standard deviation of 50. Usually the proficiency scales of the evaluations are represented as a minimum value of 0 and a maximum of 500, but, like any normal distribution, it is not limited, and it is only a convention followed by most evaluation systems in Brazil. As can be seen in Table 1, the mean value of the Portuguese language proficiency of the 9th grade students was of 246.03 and in Mathematics it was of 253.44 points.

The second group was composed of dichotomous variables, in which the reference group always assumes the value of 1. The dichotomous variables used in this study are: gender (1 if female), color (1 if white) and reprobation (1 if the student stated that they had already failed at least once). The average of these variables

indicates the proportion of participation of the group that assumes value 1. We can read these values in percentages: 52% of 9th grade students are female, 25% are white, and 33% said they have already failed once or more.

Finally, the third group of variables is the group with the standardized distribution on a scale from 0 to 10, in which the values 0 and 10 are assigned according to the limits of the indexes, where zero refers to the school (s) on which school actors demonstrated the worst perception or worst attitudes about school pedagogical practices and 10 refers to the school (s) which school actors demonstrated the best perception and attitudes about the practices, taking into account the factors evaluated through the questionnaire items. In the standardized variables 0 to 10, the concrete meaning of an additional unit in the measure is the distance of a decile between cases.

The variables of this group are the socioeconomic level of the students and the average socioeconomic level of the school, the percentage of female students in the school, the percentage of students who declared themselves white in school, the percentage of students who said they had at least one reprobation in the school and, finally, the index of pedagogical attitudes and practices - IPP. Thus, schools have an average composition of 52% of female students, 25.2% of whites and 29.4% of students who have already failed.

TABLE 1. Descriptive Statistics of Variables

Variables	Coding	N	Min.	Max.	Ave.	S.D.
Proficiency in Portuguese Language	Scale from 0 to 500	32644	97,50	381,04	246,03	46,87
Proficiency in Mathematics	Scale from 0 to 500	32644	118,36	407,71	253,44	44,79
Gender	Female = 1 and Male = 0	32644	0,00	1,00	0,52	0,50
Color/Race	White = 1 and Other = 0	32644	0,00	1,00	0,25	0,43
Socioeconomic Level	Index ranging from 0 to 10	32644	0,00	10,00	5,78	1,48
Reprobation	Already failed = 1 and has never failed = 0	32644	0,00	1,00	0,33	0,47
Girls Percentage	Scale from 0 to 10 where one unit represents 10%	811	1,54	8,00	5,17	0,59
White Percentage	Scale from 0 to 10 where one unit represents 10%	811	0,00	9,26	2,52	1,24
Average Socioeconomic Level	ISE average ranging from 0 to 10	811	2,23	5,77	4,12	0,51
Percentage Students Disapproved	Scale from 0 to 10 where one unit represents 10%	811	0,00	7,33	2,94	0,88
Index of Pedagogical Attitudes and Practices	Index ranging from 0 to 10	811	1,61	7,02	4,30	0,55

Source: Elaboration by authors based on PAEBES 2015.

The descriptive analysis allowed to know the distribution of the factors and the main characteristics of the respondents and their schools. The next step was to identify **if**, in what **direction** and **how much** each factor influences proficiency in Portuguese Language and Mathematics. Before, however, we estimate the fully unconditional model or null model, that is, a model that does not specify factors at any level. The objective of the estimation of the

null model is to verify, through the Intraclass *Correlation Coefficient* (ICC),³ if there is a significantly high variance between the evaluated schools. This coefficient can vary from 0 to 1, and the closer to zero is its value, the greater the homogeneity between the schools, being the student's school performance independent of the school in which he studies (FERRÃO, 2003). Therefore, the hierarchical regression model is only justified if there is a certain heterogeneity of performance among the schools studied. If not, the multivariate linear regression model is sufficient.

The ICC found in the null model for Portuguese Language was 0.10, that is, about 10% of the variance of student performance is associated with variability between schools. In the null model for mathematics, the ICC was 0.17, which indicates that 17% of the performance variance may be associated with variation among the study schools. Although they are not high values, the specialized literature in the field of social sciences considers them sufficient for the estimation of a hierarchical model⁴.

The hierarchical regression model was performed with variables⁵ at the student level (first level), and with variables at the school level (second level). All student-level variables are centered on the group average, and the school-level variables are centered on the large average. This ensures that the effects of variables and composition do not mix (RAUDENBUSH & BRYK, 2002). In the final model, most of the characteristics assessed at the student level and at the school level are statistically significant, that is, almost all have an influence on the explanation of the proficiencies, some in larger and others in lower value. The only variable that was not statistically significant in the final model was the percentage of girls in school, when associated with proficiency in mathematics.

As we have pointed out above, the direction a characteristic influences proficiency can be expressed by the sign of its number. Positive values mean that the factor increases performance, while negative values indicate that the factor generates a decrease in proficiency. In the final models of the present study, the reprobation generates a decrease in proficiency in Portuguese and Mathematics and the female gender generates a decrease in the performance in Mathematics. Although the data indicate that the percentage of girls has a negative effect on expected mathematics proficiency, this effect is not statistically significant, as previously mentioned (confidence interval less than 95%). The other factors have a significant impact (confidence interval greater than 95%) and positive on both proficiencies.

The first results presented refer to the hierarchical linear model determined for the analysis of the factors associated with proficiency in Portuguese Language in the 9th year of Primary Education. This two-level model was able to reduce the explanation of performance by variance when compared to the null model, that is, after the insertion of the variables in the model, 6% of Portuguese language performance continued to be explained by the variability between schools. Graph 1 shows the nine variables that made up the model and the impact of each variable on proficiency in Portuguese Language.

The value highlighted at the top of the graph refers to the intercept, that is, the expected proficiency in Portuguese when all model variables assume the value of zero. This value is 248 points. As expected, the first variable of the first level included in the model - female gender - had a positive effect on proficiency. According to the data, the expected proficiency for girls is 14 points higher than expected for boys. This finding corroborates the findings of other educational surveys that have shown that girls tend to have higher proficiency in Portuguese than boys (CITTADIN AND FRANÇA, 2016). Students who declared themselves to be white (a) would have an expected proficiency 4 points higher than students who declared themselves to be brown, black, yellow or indigenous, a result in line with the findings of Soares and Alves (2003) and by Laros, Marciano and Andrade (2010), who demonstrated an inequality in proficiency among students of color/race different from white. The third characteristic of the model refers to the socioeconomic level of the students. According to the data the better the socioeconomic condition, the greater the expected proficiency. In this sense, the result informs that a unit ahead in the scale created for the socioeconomic index, the expected proficiency increases by 4 points. This finding is in agreement with the literature as many studies find a positive correlation between the student's socioeconomic index and their performance (LARIOS, MARCIANO and ANDRADE 2010; SOARES, 2004). However, according to Soares (2004), it is expected that the average socioeconomic level of the school always has a greater effect on performance than just this variable at the individual level.

With regard to reprobation, we can see that students who experienced the reprobation experience at least once have 22 points of disadvantage in expected proficiency when compared to students who have never been disapproved. This characteristic was the one that showed the greatest (negative) impact among all of them that compose the model, corroborating the findings of other studies,

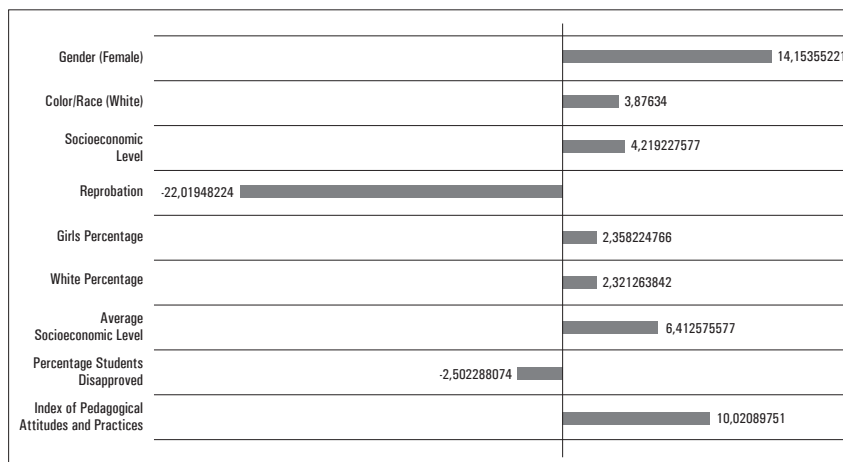
which point to reprobation as the most influential characteristic for explaining performance (ALBERNAZ, FERREIRA & FRANCO, 2002). The reprobation is a discussed theme in the educational field, and many studies in the national and international context (RIBEIRO, 1991; CRAHAY, 2006; JACOMINI, 2009; CORREA, 2013) have already shown that such a school practice – reprobation – does little to contribute to satisfactory learning. According to Crahay (2006, p.243), “repeating a year and (re)start over all schedule course does not help students in difficulty to overcome the obstacles that prevent them from being honorably successful in school.” Although the student has difficulty and did not consolidate the learning from the school year that he did, it is suggested that other strategies be implemented. Crahay (2006, 243) also emphasizes that the reprobation is not a school action that helps students in difficulty, and therefore “it is necessary to look for other means to solve this important problem.”

At the school level it was possible to control the influence of school composition on performance. The results showed that the percentage of female students in the school is responsible for increasing the expected value of proficiency by 2 points. The percentage of whites in school also increases by 2 points the expected proficiency and the average socioeconomic level is associated with an expected proficiency of more than 6 points when compared to the lowest socioeconomic level. As observed in the study by Soares (2004), the average socioeconomic level of the school has a greater effect than the socioeconomic level of the student in the performance. The percentage of failed students in school has a 2-point disadvantage when compared to the percentage of students who have never been disapproved.

Finally, the last variable of the model refers to the index of attitudes and pedagogical practices (IPP), one of the factors of effectiveness pointed out in the literature. As already mentioned, this index was formed from the perception of students and teachers about teachers’ pedagogical practices. The way the teacher administers the classroom, its methodology and interaction with the students, are directly related to the process and teaching-learning and to the performance. In this sense, as already expected, the results of the hierarchical model demonstrate that IPP significantly influences Portuguese language proficiency, increasing the student’s performance by 10 points if their teachers present attitudes and pedagogical practices evaluated as effective. This result highlights an important aspect of the school context, since it reinforces the importance of the teacher’s role in the teaching-learning process. In

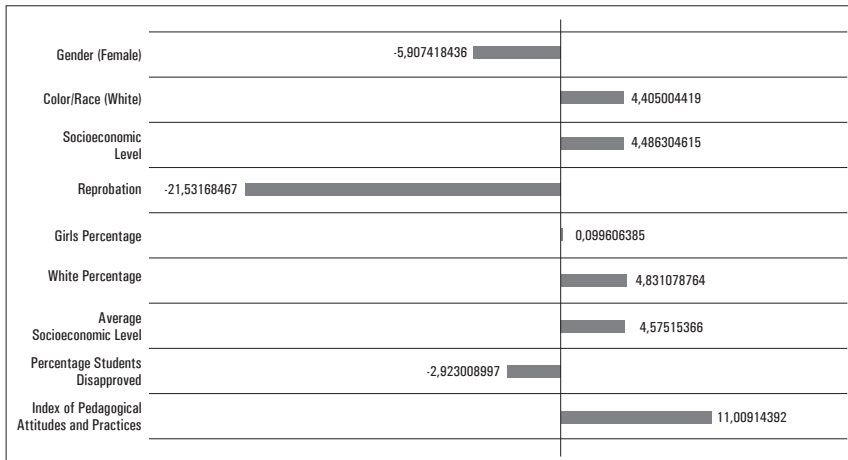
addition, it allows a critical analysis of how these attitudes are being constructed and how these practices are being developed within the school and, more specifically, within the classroom, where the teacher-student interaction occurs and where is the *locus* of learning.

GRAPHIC 1. Results of the Linear Hierarchical Model for Portuguese Language



Source: Elaboration by authors based on PAEBES 2015.

In graph 2, it is possible to verify the influence of the same factors on the expected mathematical proficiency. The value of the intercept, that is, of the expected mathematical proficiency when all the model variables assume the value of zero, is 256 points. The female, as expected, has a negative effect of 6 points on the expected proficiency in Mathematics. That data corroborates the findings of the literature, since many scholars report that girls have a greater facility in cognitive tests of Portuguese Language and boys have a greater facility in cognitive tests of Mathematics (CITTADIN AND FRANCE, 2016). Students who declared themselves to be white (a) had a 4-point increase in expected proficiency, as did students with a high socioeconomic status. Students who have been disapproved at least once have a disadvantage compared to students who have never been disapproved. Failure to do so decreases the student's expected proficiency by 22 points, an important research factor since it is repeated for the two disciplines evaluated. It is important to emphasize, therefore, the importance of research and the construction of strategies that diminish the culture of reprobation (COSTA RIBEIRO, 1991) as a pedagogical strategy and to increase learning.

GRAPH 2. Results of the Hierarchical Linear Model for Mathematics

Source: Elaboration by authors based on PAEBES 2015.

The composition of girls at school was a variable that did not present a statistically significant result, which means that the high percentage of girls in school did not significantly influence the performance in mathematics. On the other hand, the percentage of white students in school and the socioeconomic level have a positive and significant effect of 5 points in the expected proficiency. As in Portuguese language, being rejected has a negative effect of 3 points on proficiency in Mathematics. Finally, confirming the hypothesis of the study, effective pedagogical attitudes and practices within the school had an effect of 11 points on the expected mathematical proficiency. These results corroborate the data found in the literature that point out the importance of the role of the teacher in promoting cognitive development and its influence on teaching and learning processes (HANUSHEK, 1986; ALBERNAZ, FERREIRA and FRANCO; 2002, MARCONI, 2012).

However, it is necessary make a caveat about the limitation of IPP data. The results found refer to specific pedagogical practices (those included in the PAEBES contextual questionnaire) pointed out by the literature (SAMMONS, 199; OLIVEIRA, 2012; MORICONI, 2012) as practices that promote effective learning - for example: correcting homework. In this sense, the index was formulated from the perception that the school actors have about these specific pedagogical practices. There is no doubt that there are many other practices being worked in the school context, more specifically in the classroom, and all these practices are made by multiple and

complex interactions, that are not captured by the questionnaires and therefore, do not make up the index. Considering the conceptual cut about what means pedagogical practices in this study, we identified that the good perception about this restricted number of practices is positively associated with performance. Despite the conceptual limitation that underlies the construction of the IPP, the results are reliable and meet what the literature has already identified as effective practices promoting learning. This result reinforces the efficient use of these practices in the teaching and learning processes of schools.

FINAL CONSIDERATIONS

The teaching and learning processes materialize in complex realities that are permeated by numerous factors that interact and directly and indirectly affect student learning.

Considering that complex realities require complex analytical models (GOLDSTEIN, 2001), this study aimed to investigate the effect of the Pedagogical Attitudes and Practices Index (PPI) on students' performance, using a hierarchical regression model (HLM). The multilevel approach is characterized as the most suitable for investigating school factors associated with performance, since students are necessarily grouped into school units. That is, there are different levels of analysis - the student's (involving students' own characteristics such as gender and race / color) and school (which involve school characteristics, as a percentage of girls / boys and contextual factors, among which can be cited the IPP).

The main results corroborate other research findings aligned to the field of school effectiveness and reiterate the significant and positive effect that pedagogical practices considered effective in the literature on performance in both Portuguese and Mathematics. The effect of IPP on performance was controlled by socioeconomic characteristics (gender, race / color, NSE, and reprobation) - those identified as the main ones for explaining performance. In addition, the effect of IPP was observed in isolation, without considering other school aspects (school management, school climate, among others) that can also help to explain performance.

In spite of this, the effect of the IPP on the performance of both disciplines was very significant and evidences the need to look more closely at the pedagogical practices and the interactions - overlapping those involving student and teacher - that are established within the school and the classroom.

The recurrent research results evidencing the positive influence that good attitudes and pedagogic practices have on learning, should be considered by school administrators and by educational policy makers and implementers. Identifying and mapping good teacher attitudes and practices, whether through a quantitative, qualitative or mixed approach, can provide important subsidies for initial and continuing training actions, as well as can be used as a subsidy for teachers and other school agents to rethink their attitudes and pedagogical practices in order to guarantee student learning.

REFERENCES

ALBERNAZ, A.; FERREIRA, F.; FRANCO, C. Qualidade e equidade no ensino fundamental brasileiro. **Pesquisa e Planejamento Econômico**, v.32, n.3, pp. 453-476, 2002.

BARROS, R. P., MENDONÇA, R., SANTOS, D. D., QUINTAES, G. Determinantes do desempenho educacional no Brasil. **Pesquisa e Planejamento Econômico**, n.31, n.1, pp. 1-42, 2001.

BRESSOUX, P. As pesquisas sobre efeito-escola e o efeito-professor. In: *Educação em Revista*, n. 38, 2003.

BROOKE, N., SOARES, J.F. (Orgs.). **Pesquisa em eficácia escolar: origem e trajetória**. Editora UFMG, 2008.

BROOKE, N. Eficácia escolar. In: OLIVEIRA, D.A.; DUARTE, A.M.C.; VIEIRA, L.M.F. **DICIONÁRIO: trabalho, profissão e condição docente**. Belo Horizonte: UFMG/Faculdade de Educação, 2010. CDROM.

CARDELLI, D.T., ELLIOT, L.G. **Avaliação por diferentes olhares: fatores que explicam o sucesso de escola carioca em área de risco**. Ensaio: aval. pol. públ. Educ., Rio de Janeiro, v. 20, n. 77, p. 769-798, out./dez. 2012.

CITTADIN, I., FRANÇA, M.T.A. A violência como fator de influência no desempenho de estudantes do ensino fundamental: análise utilizando modelos multiníveis. In: **XIX Encontro de Economia da Região Sul**, 2016. Disponível em: <https://www.anpec.org.br/sul/2016/submissao/files_I/i29ec27d4bf42cdc1ad3a106dd7e3dcd5f.pdf>. Acesso em: 24 de abril de 2018.

CLOTFELTER, C., LADD, H., VIGDOR, J. Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. **Economics of Education Review**, v. 26 n. 6, pp. 673-682, 2007.

CORREA, E. **Efeito da repetência nos anos iniciais do Ensino Fundamental: um estudo longitudinal a partir do Geres**. 2013, 122 fls. Dissertação (Mestrado em Educação) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013.

COSTA RIBEIRO, S. A pedagogia da repetência. **Estudos avançados**, v.5, n.12., 1991.

CRAHAY, M. É possível tirar conclusões sobre os efeitos da repetência? **Cadernos de Pesquisa**, v. 36, n. 127, p. 223-246, jan./abr. 2006.

_____. **Poderá a escola ser justa e eficaz?** Instituto Piaget, Lisboa, 2002.

DALMORO, M.; VIEIRA, K. Dilemas na construção de escalas tipo Likert: o número de itens e a disposição influenciam nos resultados? **Revista Gestão Organizacional**, v.6 – Edição Especial, Chapecó, 2013.

DARLING-HAMMOND, L., HOLTZMAN, D. J., GATLIN, S. J., HEILIG, J. V. Does teacher preparation matter? Evidence about Teacher Certification, Teach for America, and Teacher Effectiveness. **Education Policy Analysis Archives**, v.13, n.42, pp. 1-51, 2005.

FERRÃO, M. E. **Introdução aos modelos de regressão multinível em educação**. Campinas: Editora Komedi, 2003.

FIELD, A. **Descobrimo a estatística usando o SPSS**. Porto Alegre: Artmed, 2009.

GAUTHIER, C. MELLOUKI, M. SIMARD, D. BISSONNETTE, S. RICHARD, M. **Quelles sont les pédagogies efficaces? Un état de la recherche**. Fondation pour l'innovation politique, 2005. Documento disponível em: <<http://www.fondapol.org/>>. Acesso em: 12 de dezembro de 2017.

GOLDSTEIN, H. Modelo de realidade: novas abordagens para a compreensão de processos educacionais. In: Franco, C. **Avaliação, ciclos e promoção na Educação**. Porto Alegre: Artmed, 2001.

HANUSHEK, E.A. The economics of schooling: production and efficiency in public schools. **Journal of Economic Literature**, v.24, n.3, pp. 1141-1177, 1986.

HANUSHEK, E. A., RIVKIN, S.G. How to improve the supply of high-quality teachers. **Brookings Papers on Education Policy**, n. 7, p.7-25, 2004.

JACOMINI, M. Educar sem reprovar: desafio de uma escola para todos. **Educação e Pesquisa**, São Paulo, v.35, n.3, p. 557-572, set./dez. 2009.

LAROS, J.A., MARCIANO, J.L.P., ANDRADE, J.M. **Fatores que afetam o desempenho na prova de Matemática do SAEB: um estudo multinível**. Avaliação Psicológica, 2010, 9(2), pp. 173-186, 2010.

MACHADO SOARES, T. Influência do professor e do ambiente em sala de aula sobre a proficiência alcançada pelos estudantes avaliados pelo Simave 2002. **Estudos em Avaliação Educacional**, v.28, pp.103-123, 2003.

MORICONI, G.M. **Medindo a eficácia dos professores: o uso de modelos agregados para estimar o efeito do professor**. Tese de Doutorado, Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas, 2012. Disponível em: <<https://bibliotecadigital.fgv.br/dspace/handle/10438/9620>>. Acesso em: 29 de novembro de 2017.

MORTIMORE, P. **The nature and findings of school effectiveness research in the primary sector**. In: RIDDELL, S.; BROWN, S. (Ed.). School effectiveness research: its messages for school improvement. London: HMSO, 1991.

MURAKI, E.; BOCK, R. D. **PARSCALE: IRT based test scoring and item analysis for graded opened exercises and performance task**. Version 3. Chicago: Scientific Software, (1996).

NYE, B., KONSTANTOPOULOS, S., HEDGES, L. V. How large are teachers effects? **Educational Evaluation and Policy Analysis**, v.26, n.3, pp.237-257, 2004.

OECD, **PISA 2012 Results: Ready to Learn (Volume III): Students' Engagement, Drive and Self-Beliefs**, PISA, OECD Publishing, Paris, 2013. <https://doi.org/10.1787/9789264201170-en>. Acesso em: 29 de novembro de 2017.

OLIVEIRA, L.H.G. **Habilidades de leitura e práticas pedagógicas associadas ao seu aprendizado**. Tese de Doutorado, Departamento de Educação da Pontifícia Universidade Católica do Rio de Janeiro, 2012. Disponível em: <http://www2.dbd.puc-rio.br/pergamum/biblioteca/php/mostrateses.php?open=1&arqtse=0811296_2012_Indice.html>. Acesso em: 20 de dezembro de 2017.

RAUDENBUSH, S.W. & BRYK, A.S. **Hierarchical Linear Models: Applications and Data Analysis Methods**. Thousand Oaks: SAGE, 2002.

RIBEIRO, Sérgio da Costa. A pedagogia da repetência. **Estudos Avançados**, São Paulo, USP, vol. 5, nº 12, mai.-ago. de 1991.

RIVKIN, S. G., HANUSHEK, E. A., & KAIN, J. F. Teachers, schools, and academic achievement. **Econometrica**, v.73, n.2, pp.417-458, 2005.

ROCKOFF, J. E. The impact of individual teachers on student achievement: evidence from panel data. **The American Economic Review**, v.94, n.2, pp. 247-252, 2004.

SAMEJIMA, F. A. **Estimation of latent ability using a response pattern of graded scores**. Psychometric Monograph, 17, (1969).

SAMMONS, P. **School Effectiveness: cominf of age in the twenty-first century**. Lisse: Swets & Zeitlinger, 1999.

SIEGEL, S.; CASTELLAN Jr., J.N. **Estatística não-paramétrica para ciências do comportamento**. São Paulo: Artmed, 2006.

SOARES, J. F. O efeito da escola no desempenho cognitivo de seus estudantes. **Revista Electrónica Iberoamericana sobre Calidad, Eficácia y Cambio en Educación**, 2(2), 83-104, 2004.

SOARES, J. F. Qualidade e equidade na educação básica brasileira: fatos e possibilidades. In Brock, C., & Schwartzman, S., **Os Desafios da Educação no Brasil**. Nova Fronteira, pp. 91-117, 2005.

SOARES, J. F., ALVES, M.T.G. Desigualdades raciais no sistema brasileiro de educação básica. **Educação e Pesquisa**, 29(1), 147-165, 2003.

SOARES, J.F., ALVES, M.T.G. Efeito-escola e estratificação escolar: o impacto da composição de turmas por nível de habilidade dos estudantes. **Educação em Revista**, v.45, pp.25-58, 2007.

WRIGHT, S.P., SANDERS, W.L., HORN, S.P. Teacher and classroom context effects on student achievement: implications for teacher evaluation. **Journal of personnel evaluation in education**, v.11, p.57-67, 1997.

NOTES

¹ Proficiencies are calculated on the basis of Item Response Theory (IRT).

² The items were constructed based on the literature on school effectiveness and teacher effectiveness and reflect theoretical definitions about good pedagogical practice.

³ ICC calculation: $\rho = \tau_{00} / (\tau_{00} + \sigma^2)$

⁴ Full statistics of the models can be found in the appendix.

⁵ For all variables we found statistically significant values, considering a confidence level of 95%.

Submission: 07/05/2018

Approbation: 07/09/2018

Contact:

Rua Marquês de São Vicente, n. 225, 10º andar,
Sala 1049, Prédio Cardeal Leme, Gávea
Rio de Janeiro | RJ | Brasil
CEP 22.793-260

APPENDIX

Table 2. Estimates of Fixed Effects (Portuguese Language) - Null Model

Parameters	Estimate	Standard Error	df	t	Significance	95% confidence interval	
						Lower limit	Upper limit
Intercept	245.763	0.624	744.546	393.884	0.000	244.538	246.988

TABLE 3. Estimates of Covariance Parameters (Portuguese Language) - Null Model

Parameters	Estimate	Standard Error	Wald Z	Significance	95% confidence interval		
					Upper limit	Lower limit	
Waste	1964.433	15.462	127.049	0.000	1934.361	1994.973	
School Intercept	Variance	230.211	15.137	15.209	0.000	202.376	261.874

TABLE 4. Estimates of Fixed Effects (Mathematics) - Null Model

Parameters	Estimate	Standard Error	df	t	Significance	95% confidence interval	
						Lower limit	Upper limit
Intercept	254.470	0.729	736.346	349.078	0.000	253.039	255.902

TABLE 5. Estimates of Covariance Parameters (Mathematics) - Null Model

Parameters	Estimate	Standard Error	Wald Z	Significance	95% confidence interval		
					Upper limit	Lower limit	
Waste	1682.107	13.247	126.984	0.000	1656.343	1708.271	
School Intercept	Variance	347.140	21.131	16.428	0.000	308.099	391.128

TABLE 6. Estimates of Fixed Effects (Portuguese Language) - Two-Level Model

Parameters	Estimate	Standard Error	df	t	Significance	95% confidence interval	
						Lower limit	Upper limit
Intercept	247.927	0.511	656.260	485.240	0.000	246.924	248.930
Female	14.154	0.471	32048.755	30.045	0.000	13.230	15.077
White Color	3.876	0.564	32064.492	6.875	0.000	2.771	4.981
Class Social	4.219	0.176	32124.839	23.907	0.000	3.873	4.565
Reprobation	-22.019	0.514	32146.937	-42.851	0.000	-23.027	-21.012
Percentage of Girls	2.358	0.705	973.346	3.343	0.001	0.974	3.743
White Percentage Average Socioeconomic Level	2.321	0.374	857.454	6.206	0.000	1.587	3.055
Percentage Students Disapproved	6.413	0.624	772.250	10.270	0.000	5.187	7.638
Index of Pedagogical Attitudes and Practices	-2.502	0.549	813.982	-4.557	0.000	-3.580	-1.424
	10.021	0.893	793.799	11.227	0.000	8,269	11.773

TABLE 7. Estimates of Covariance Parameters (Portuguese Language) - Two-Level Model

Parameters	Estimate	Standard Error	Wald Z	Significance	95% confidence interval		
					Upper limit	Lower limit	
Waste	1738.896	13.788	126.117	0.000	1712.081	1766.131	
School Intercept	Variance	127.105	9.479	13.409	0.000	109.820	147.110

TABLE 8. Estimates of Fixed Effects (Mathematics) - Two-Level Model

Parameters	Estimate	Standard Error	df	t	Significance	95% confidence interval	
						Lower limit	Upper limit
Intercept	255.948	0.588	637.588	434.961	0.000	254.792	257.103
Female	-5.907	0.440	31936.659	-13.426	0.000	-6.770	-5.045
White Color	4.405	0.527	31954.010	8.364	0.000	3.373	5.437
Class Social	4.486	0.165	31991.960	27.210	0.000	4.163	4.809
Reprobation	-21.532	0.480	32015.747	-44.847	0.000	-22.473	-20.591
Percentage of Girls	0.100	0.785	874.764	0.127	0.899	-1.441	1,641
White Percentage	4.831	421	781.627	11.479	0.000	4.005	5.657
Average Socioeconomic Level	4.575	0.708	729.037	6.458	0.000	3.184	5.966
Percentage Students Disapproved	-2.923	0.620	761.143	-4.713	0.000	-4.140	-1.706
Index of Pedagogical Attitudes and Practices	11.009	1.010	750.351	10.904	0.000	9.027	12.991

Table 9. Estimates of Covariance Parameters (Mathematics) - Two-Level Model

Parameters	Estimate	Standard Error	Wald Z	Significance	95% confidence interval	
					Upper limit	Lower limit
Waste	1514.993	12.025	125.987	0.000	1491.606	1538.745
School Intercept	189.490	12.943	14.640	0.000	165.745	216.634