



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

Factors associated with the academic performance of 3rd grade students in Mozambique: An application of the multilevel linear model.

Factors associated with the academic performance of 3rd grade students in Mozambique: An application of the multilevel linear model

I admire Nacuata – Rovuma University

Abubacar Nizar – Rovuma University

Abstract

Introduction: Academic performance is influenced by a combination of individual, family, school, and contextual factors. Understanding these factors is essential for guiding effective educational policies. **Objective:** To determine the factors influencing the academic performance of 3rd- grade students in Mozambique, based on empirical evidence. **Methodology:** A quantitative, cross-sectional study with a sample of 7735 students randomly selected using cluster sampling. A proficiency test, a teacher's notebook, a school administration questionnaire, and a school newsletter were administered to the students. The analysis includes descriptive statistics and multilevel linear regression, all performed using SPSS. **Results:** 45% of the variance in the performance of 3rd-grade students is attributed to differences between schools. In the final model, the variables age and frequency with which the student speaks Portuguese outside of school, labeled in the database as "Mother Tongue," were significant at the 5% significance level. At the school level, no variable was significant. **Conclusion:** The age and frequency with which students speak Portuguese outside of school (as their mother tongue) are factors that affect students' academic performance, thus highlighting the need to create community-based reading education programs, because, according to many authors, this is where the biggest problem lies.

Keywords: Academic performance, level, variance.

Abstract:

Introduction: Academic performance is influenced by a combination of individual, family, school, and contextual factors. Understanding these factors is essential to guide effective educational policies.

Objective: To determine the factors that affect the academic performance of 3rd grade students in Mozambique, based on empirical evidence. **Methodology:** A quantitative, cross-sectional study was conducted with a random sample of 7735 students, using cluster sampling. A proficiency test was applied to the students, along with a teacher's notebook, a school administration questionnaire, and a school information bulletin. The analysis included descriptive statistics and multilevel linear regression, all performed using SPSS. **Results:** 45% of the variance in 3rd grade student performance is attributed to differences between the schools. In the final model, variables such as age and the frequency with which the student speaks Portuguese outside school – recoded in the database as "Mother Tongue" – were significant at the 5% level. At the school level, no variable was significant.

Conclusion: Age and mother tongue are the key factors affecting student's academic performance.

This highlights the need for the creation of community reading programs, since many authors argue that this is where the main problem lies.

Keywords: Academic performance, level, variance.

INTRODUCTION

In Mozambique, school education is generally perceived as being in crisis, which The existence of a "foreign" school, that is, a school that does not reflect reality, contributes to this. socio-cultural context of the communities in which it is embedded, to the point that some authors state that "the Schools in Mozambique are still not able to be spaces for the construction and systematization of knowledge that takes into account different anthropological, political, social and



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026
"cultural".

According to the United Nations Development Agency (USAID, 2021), although Mozambique has continued to educate an increasing number of children in schools Primary survey results indicate that obstacles to achieving the Educational objectives. This can be verified by the results of the 2016 National Assessment from the Ministry of Education and Human Development (MINEDH) which showed that only 4.9% Of the third-grade students, 6.3% reached the required level in reading and writing. recorded in the previous 2013 study (USAID, 2021). This is one of the strategic objectives of the Plan. Mozambique's Strategic Education Plan 2020-2029 aims to achieve the completion of primary education. Quality and inclusion are achieved by setting a goal of at least 59% of children completing schooling. primary education until 2029 (Ministry of Education and Human Development [MINEDH], 2022). However To achieve this goal, the political process must be supported by solid evidence regarding the factors. drivers of educational outcomes at the family, school, and community levels. Thus, one of One of the initiatives of the Mozambican government was the creation of the National Evaluation, an instrument that serves to subsidize and promote policies aimed at improving the quality of Mozambican education. through its results.

In this context, the present study was developed with the objective of analyzing the data from National Assessment 2016 relating to the subject L1 (Portuguese Language), in order to identify the Factors that affect the academic performance of 3rd grade students. However, to achieve this The researcher must remember that the student's academic performance does not depend solely on the objective. from the student himself, as he may be influenced by other factors that are grouped together. at hierarchical levels, such as the school the student attends and the city they belong to. the school.

Therefore, to analyze data arranged in a hierarchical structure, it is necessary to have... specialized techniques for this purpose, and a statistical methodology that can be used in these One of the conditions is multilevel analysis. In this type of analysis, the results are obtained by considering the... levels at which the variables are embedded (Kreft & De Leeuw, 1998; Natis, 2001). Thus, in addition with the objective of identifying the factors that influence students' academic performance through In addition to the multilevel model, this article also aims to show the potential of this class of Models for data processing in a hierarchical structure compared to classical models. For To achieve the objective of this study, the students in the 3rd grade who underwent the Assessment were considered. National Assessment (ANA) for the year 2016. This assessment was implemented by the Ministry of Education and Ministry of Human Development (MINEDH) through the National Institute of Education (INDE), starting from the year 2013 and its objective is to provide data for the indicator of the Strategic Education Plan. (PEE) and the Government's Five-Year Program (PQG) (MINEDH, 2016).

THEORETICAL FRAMEWORK

1.1. Academic Performance

Academic performance refers to the outcome of students' academic skills. when assessed in different fields of learning (Fontes, 2017).

Authors such as Heck *et al.* (2014) describe that the expression performance is used for to convey the idea of "*achievement*," that is, the act of conquering something, of being successful, through of effort and skill. According to Vernier, Bagolin and Jacinto (2015) the educational issue It is not based solely on the quantity of years of study, but primarily on the quality of those years. existing levels of education. For these authors, the most general way to assess school performance is through an educational production function, initially presented by Coleman (1966), in which seeks to explain performance based on the personal and socio-economic aspects of the students and of school supplies:

$$Y = f(A, P, D, E, \tilde{y})$$

Where Y is student performance measured by grades, and A is a feature vector. of the students, such as race, gender, and age, P is the vector of teacher characteristics, D is the vector of characteristics of the director, E contains information about the school's infrastructure, and \tilde{y} is the term for random error.

2.1.1. Evaluation of academic performance

Academic performance is evaluated by measuring intelligence, with Factors such as memory capacity, attention, reasoning, evaluation, and problem-solving. (Lemos, Almeida, Guisande, & Primi, 2008 apud Fontes, 2017). According to Haydt (1997), the evaluation the assessment of academic performance is carried out with two main objectives: (1) to identify the difficulties of (1) to help students overcome learning challenges; and (2) to evaluate the effectiveness of teaching, being considered a parameter for analyzing the work developed in the classroom and at school, thus reflecting the quality of teaching. In this context, to outline the necessary actions for the Improving the quality of the teaching-learning process requires more than just evaluating performance as... It's not just an isolated point; rather, one must understand the school context and the socio-economic factors that... However, grades and tests do not provide such information (Américo; Lacruz, 2017; Costa, 2005; Faria; Guimarães, 2015).

2.1.2. Factors that influence academic performance



2.1.2.1. Individual and family factors

The variables related to student level that positively affect academic performance, According to Laros *et al.* (2010), these are: comparison of the student with peers, with a differentiated effect of school to school; cultural resources available to the student at home, such as access to computers. with internet, books, general information magazines, newspapers, etc.; the student enjoys studying the subject, with varying effects from school to school, and the student does homework. On the other hand, the variables According to the author, the factors affecting a student's performance that negatively impact their level of ability were: the relationship The relationship between the student's family and the school, and whether the student works in addition to studying.

In the work of Menezes-Filho (2007), other variables, at the student level, had the effects. which could be expected a priori. For example, students who live with their parents (or with at least One of them performs better; those who work from home for more than four hours.

those who read books or newspapers tend to perform worse; those who read books or newspapers tend to perform better, and those who do not perform well perform worse. Those who work outside the home perform worse. Having one or more computers and more than 20 Having books at home improves learning, as does having electricity and living in small families. (with up to five people in total).

Barros *et al.* (2001), cited by Macedo (2004), investigated some of the impacts of important types of determinants of educational performance. Regarding the influence of family on educational background of children, the results of their study revealed that the schooling of parents, And in particular, the mother's [child's] is the most important variable in determining educational performance. of children. The factor of "doing homework" always has a strong and positive effect on performance. of the student, as well as the findings in Laros *et al.* (2010).

Another important variable is the age of entry into the school system: students who have Preschool children perform better in all grades compared to those who started in 1st grade. series (Menezes-Filho, 2007). Regarding social and economic factors, Gomes (2018) and Ferrão (2003), They claim that when the financial or economic conditions of families do not allow for greater Lack of care or attention to a child can lead to poor academic performance due to a lack of resources. provide them with a better quality of life. Gomes (2018) also states that behaviors Inappropriate behaviors on the part of parents or guardians, especially promiscuity, prostitution, and drugs. In the family, domestic violence, unemployment, and family malnutrition are factors that interfere. directly impacting the child's behavior, contributing to learning difficulties.

2.1.2.2. School factors

According to Menezes-Filho (2007), school-level variables, such as the number of Computers in schools, the selection process for principals and students, education, age, and salary. The influence of teachers has very little effect on student performance. According to this author, in



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

In the public system, class size does not appear to be a significant factor in explaining student performance.

However, the number of hours per class has a positive and statistically significant effect, that is, the

Students who spend more than four hours in the classroom perform better than those who...

They stay for less time than that.

Soares (2004) adds that a second factor affecting learning, with regard to

Student concentration refers to the size of the school, measured in the number of students served. According to him.

Student performance is higher in smaller schools. There is also much discussion regarding this.

digital inclusion, that is, the need to put computers in public schools. In this matter,

The results in Menezes-Filho (2007) corroborate the results of Soares (2004). According to the

First author, regarding the characteristics of the teacher, the results are quite interesting. In

First, he studied the teacher's age, which positively affects student performance.

Their studies indicated that teachers over 49 years old who remain teaching,

They are able to transmit more knowledge to their students. Secondly, the following was studied...

The teacher's level of education had a small effect on student performance.

On the contrary, he further noted that neither the time spent at the school, nor the fact of teaching in more than one

School affects student performance. Considering the findings in Luz (2006), the physical structure

The school's performance showed a significant and positive coefficient. Thus, it is possible to say that the performance

The student's performance increases as the structure of the school they attend improves, because resources grow.

available for its development and comfort in the environment where it takes place.

2.2. Multilevel Linear Model

Multilevel models, also called "random coefficient models,"

Variance component models, mixed models, and hierarchical linear models are models that

They acknowledge the existence of multilevel or hierarchical structure in the data (Raudenbush and Bryk, 2002).

They are described as models in which the type of regression analysis considers only one factor.

structure, data organized at different levels of aggregation. This organization of the data,

Designated as a hierarchical structure, it is described by the grouping of units according to certain...

characteristics that differentiate them, assigning them to different lower-level groups. These, therefore...

In turn, they belong to higher-level units, considering their common aspects, and

and so on, for example, students within classes or patients within hospitals.

individuals within their national units, individuals within organizations, members

family members within families and respondents within interviewers (Hox, 1995).

These models are seen as an evolution of the classical regression model in which...

Variables are organized and analyzed at multiple levels, thus making the estimation more accurate.

of the values for standard errors, confidence intervals, and hypothesis testing (Laros and Marciano, 2008).

According to Ferrão (2003), in classical regression models, the value of the constant and the



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

Coefficients that allow for model adjustment are fixed parameters, whereas in multilevel models

These parameters are considered random because they are influenced by the levels.

higher hierarchical levels. According to Hox (1995), in classical regression models, the tests

Statisticians rely on the assumption of independence of observations, and if this assumption is

violated the estimates of the standard errors of the conventional statistical tests are very small, and

This results in many spuriously significant outcomes.

2.2.1. Multilevel structure in the educational system According

to Goldstein (1992); Goldstein and Cuttance, (1988); Nuttall et al. (1989);

Woodhouse and Goldstein (1989) and Plewis (1991), cited by Twisk (2006), studies with analysis

Hierarchical data structures initially emerged in the field of education. Their use is

precisely related to a study conducted in the 1970s by Bennett (1976) with children

from basic education, in England, which applied linear regression models to demonstrate that students

students from a primary school in the United Kingdom subjected to the so-called formal method of teaching presented

greater progress than students who were not subjected to this teaching method, the so-called

progressive method.

Later, Aitkin et al. (1981) demonstrated that when the analysis considered the

By grouping students into classes, this difference disappeared, and students were subject to formal education.

They no longer showed significant progress. It was then that, alongside Longford (1986), they wrote a

This article revolutionized the world of educational research: *multilevel models*. In this article, they...

They demonstrated that linear regression models used to study how a set of

Variables that explain a response variable could only be used when the observations

were independent.

However, in the educational context, where students are grouped into classes, different

Classes are grouped into schools, and the different schools are grouped into municipalities or regions; it's not compatible.

with the imposition of independence of observations. This can happen according to Ferrão and

Fernandes (2000) for socio-geographic and economic or other reasons, which cause the

The distribution of students among schools should not be random, thus contributing to ensuring that students from

Students from the same school show greater similarities to each other than students from different schools.

3. MATERIALS AND METHODS

3.1. Sample

This study used data from the 2016 National Assessment (ANA).

originating from the Data Analysis Center, C. Data, of the Faculty of Research at the University

Pedagogical University (UP) Maputo. The scope of the ANA 2016 sampling includes data from 3rd grade students.

class, by provinces throughout the country. The research universe covers all students in the 3rd grade.



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

enrolled in 2016 in urban, peri-urban and rural schools, public, community and private.

For this study, students who had at least one missing piece of data were not considered, leaving

A sample comprised 7735 students from 270 schools. Of the sample considered, 51.8%

48.9% are male and 48.9% are female. Regarding the frequency with which...

Students speak Portuguese outside of school; 51.8% stated that they had never spoken Portuguese outside of school.

from school, followed by the 22.5% who spoke sometimes. Those who spoke often and always

They constitute a minority of the sample, with 12.3% for the "often" category and 13.3% for the "often" category.

always.

3.2. Instruments In

addition to the L1 (Portuguese Language) test, with 24 questions answered by the students,

The following instruments were also used:

– *Teacher's Notebook*: the teachers answered a questionnaire composed of questions

Regarding professional training and experience, curriculum and teaching practices, and school conditions;

– *Director's Questionnaire*: a questionnaire composed of questions about training and

professional experience, curriculum and teaching practices, management methods and school conditions

It was answered by each school principal;

– *School Newsletter*: a document composed of questions about aspects

related to the school's infrastructure, environment, and overall conditions, and answered by

deputy director.

3.3. Data analysis procedures

First, the data were analyzed using descriptive measures, in a way that...

stratified by region, to understand the behavior of academic performance over time.

regions. Next, a multilevel regression model was estimated to explain the performance.

academic performance of the students. First, the model with only the intercept (the null model) was estimated, which,

According to De Jesus and Laros (2004); Heck *et al.* (2014); Hox (2002), it serves as a basis for comparison.

with subsequent models. In a second phase, the explanatory variables of were included.

Level 1 with a fixed effect. Subsequently, the explanatory variables for level 2 were inserted. And by

Finally, it was assessed whether any of the slope coefficients of the level 1 explanatory variables

It had a significant variance component between schools, that is, it was verified whether the variables

The student's performance showed different effects across schools.

4. RESULTS AND DISCUSSION

4.1. Academic Performance The

distribution of grades at the national level centers around 10.93 points, with a

Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

A deviation of 3.84 points. The Southern region had the best academic performance, being represented by...

Inhambane province, with an average score of 12.99, followed by the Central region,

represented by Sofala, with an average value of 12.52. The northern zone continues to have very poor results.

academic performance results corroborate David's study (2017, p. 6), which states

that "the northern provinces show the worst results in social education indicators (...)

and Timbane (2014) who states that the northern provinces of Mozambique still have high rates

of illiteracy, a fact that is justified by the distribution of the population across remote areas and of

Uneven investment in infrastructure by the Government.

By correlating these two findings, it makes sense that the high illiteracy rate of

The population can have a negative impact on students' academic performance, since...

An illiterate individual has limited capacity from an academic perspective; they are unable to...

to recognize the value of formal education, and often sends their children to school as a

It's a simple social rule, which means he's not a present father during his children's school years.

The results described above are shown in Table 1.

Table 1. Description of academic performance

Region	Province	n	Minimum	Maximum	Average	SD
North	Niassa	275	1.67	15.83	8.16	2.43
	Cabo Delgado	517	2.50	16.67	9.49	2.98
	Nampula	1144	0.83	18.33	9.53	3.68
Center	Zambezia	737	0.00	19.17	11.04	3.94
	Tete	803	0.00	19.17	10.72	3.94
	Manica	836	2.50	18.33	9.72	3.83
	Sofala	344	0.83	19.17	12.52	3.98
South	Inhambane	845	3.33	19.17	12.99	3.50
	Gaza	929	1.67	19.17	10.93	3.29
	Maputo Province	918	3.33	20.00	12.31	3.47
	Maputo City	387	3.33	20.00	12.64	3.64
National		7735	0.00	20.00	10.93	3.84

4.2. Estimation of the Academic Performance Model

For model estimation, a multilevel approach was used, applying it to the data.

Educational data from the 2016 National Assessment. The data in this study were collected from students...

from the 3rd grade to whom the literacy proficiency test was applied.

The grade obtained by the students in the assessment was used as the dependent variable (performance).

academic), measured on a scale of 0 to 20. The explanatory variables of the student level, which

The factors considered were *gender, age, and native language (frequency with which the student speaks).*

Portuguese outside of school). The explanatory variables considered, associated with school, are the *nature*

of the school, the school's location, the type of school, the teacher's age, the academic level of the

professor, the professor's years of experience, the professor's workload, and the ratio

teacher/student.

In estimating the model that characterizes students' academic performance, generally

Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

One begins by estimating the model without any explanatory variables, in order to obtain information.

Regarding the amount of observed variance in performance within each school, that is, at level 1 (grouping) and between schools (i.e., at level 2). This model provides an estimate of the overall average across all students and schools. Another relevant piece of information that can be obtained from this model is regarding the existence of evidence of clustering structure in the data, component testing of variance of level 1 and variance of intercepts along the units of level 2, if this exists evidence can introduce bias in parameter estimates and standard errors, leading to inferences. Incorrect, using a classic model.

Table 2 presents the significance test of the intercept or mean value of the classification.

in literacy across all schools. The hypotheses under test are: $H_0: \bar{y}_{00} = 0$ vs $H_1: \bar{y}_{00} \neq 0$. Being $F(1, 266.291) = 4003.282$, $p < 0.05$, H_0 is rejected and it is concluded that the overall mean for literacy is significantly greater than zero. The estimate of this average can be found in Table 3, page 8 ($\bar{y}_{00} = 10.48421$). Since $p < 0.05$, it can also be concluded that the estimate remains significantly greater than zero.

Table 2. Fixed Effects Test of the intercept

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	266,291	4003,282,000	

On page 8, the results of the model estimation can be seen in Table 3.

null (Model 1). The estimate of the variance of the residuals is $\bar{y}_{\text{Hey}}^2 = 8.373$. Since $p < 0.05$, it is rejected. H_0 and it is concluded that the residual variance is significantly greater than zero. It is possible to calculate the ratio of the variance of the residuals to the overall mean ($8.373/10.484$), which is equal to approximately 80% of this (global average). Thus, it can be said that there is a large quantity of variation in the results of the dependent variable not explained by the model. This shows that there is a large quantity of factors contributing to the occurrence of the phenomenon that were not considered in the model. This also means that within schools, students show very dispersed individual results.

of the school average. Similarly, the estimate of the variance of the random term at the highest level.

$u_{0j}^2 = 6.894$ and since $p < 0.05$, we can reject H_0 and conclude that the variance between schools is \bar{y}

significantly greater than zero. Calculating the proportion of the variance of the intercepts with respect to the global average ($6,894/10,484$) yields a variance of approximately 66% of this. This result demonstrates that schools differ significantly in their average performance rankings. This indicates the specific characteristics of each school. They have a significant impact on results, which implies the need for differentiated strategies. for different schools, since the characteristics that influence student performance can vary significantly between schools.



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

To further emphasize this conclusion, based on the information regarding the variance of For levels 1 and 2, an intra-class correlation of 0.45 was calculated, that is, $6.894 / (8.373 + 6.894)$, which suggests that, initially, 45% of the variation in students' academic performance results is explained by the differences between schools, which makes it necessary to use an approach multilevel. Therefore, as Maroco (2011) states, the non-consideration of the level 2 factor (schools) and the lack of independence of its elements would result in statistical tests with a higher probability to commit a Type I error (finding parameters that are not actually significant).

In this model, the estimate of the χ^2 LLREML (deviance) function is equal to 39186.593. The REML Restricted Maximum Likelihood (RML) is the estimation method of Standard parameters in SPSS Statistics. Therefore, if you intend to compare models, you should use the method. FML (Full Maximum Likelihood). The value obtained from $-2LL$ This model serves as a benchmark for comparison to verify the fit of subsequent models.

According to Snijders and Bosker (1999), deviance reflects the lack of fit between the data and the The model, not being interpreted on its own, but rather by the magnitude of the differences between its value and that found in the following model. In this article, the contribution of will be considered significant.

a variable when the difference between the deviances of two models can be called *chi-*

The squared change is larger when compared to the χ^2 statistic. ² with degrees of freedom

corresponding to the difference in the number of parameters of the current model compared to the previous model (Field, 2013, p. 957) or when the difference between the deviance of the two models is divided by the degrees of freedom, that is, the number of parameters added to a model compared to the previous one, for greater than 1.96, which can approach 2 (De Jesus & Laros, 2004, p. 101).

After the relevance of using the multilevel approach was proven, the next step was... included the low-level (student) variables capable of explaining its variance. The following were included...

The following variables: *gender* (coded 1 – Male and 2 – Female), *age*, and *native language*.

(*frequency with which the student speaks Portuguese outside of school*), coded 1 – Never, 2 – As

1 – times, 3 – Often, and 4 – Always. The variables *gender* and *mother tongue* were considered.

as factors in SPSS because they were not transformed into *dummy variables*. This fact served to

to show that it is possible to work in the multilevel model with categorical variables that are not

Dummy variables, because SPSS recognizes them as factors, transforms them into *dummy variables* and

It automatically assigns them a reference category.

The first question at this stage is to verify whether the inclusion of level 1 variables caused or not an improvement on the model. Therefore, since the objective was to compare the models, the method used The FML (Functional Multiplication Method) was used to estimate the parameters. Thus, comparing the values of the information criteria...

One can select the model that best fits the data, such as the one with the lowest value. In this case...

In this sense, the current model is the one that presents the best fit in relation to the model without any variables.



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

Explanatory (Model 1) [(D2 = 38777.755 < D1 = 39186.593)]. Once the improvement was observed

From the model, it is now important to test whether this improvement is statistically significant or not, and for that...

The difference in deviance (\bar{y}_{2LL}) between the two models was used, as illustrated by the following equation.

$$(\bar{y}_{\text{Change}})^2 = 39186.593 - 38777.755 = 408.838; \text{GLMudança} = 8 - 3 = 5).$$

The critical value of the chi-square statistic with five degrees of freedom, at the level of

The significance level of 5% is 11.07. Therefore, comparing the chi-square value of change with the value

The critical chi-square test concludes that the inclusion of the student variables (Level 1) created a change.

significant [$D > \bar{y}_w^2 (5)$]. Another applicable procedure consists of finding the ratio between the difference

of deviance and degrees of freedom, where if the ratio is greater than two, the contribution will be

considered significant. In this case, we have $408,838 / 5 = 81,768$, which is greater than 2.

The results of Model 2 can be seen in the first column (Model 2) of the Table.

3, which can be found on page 8. It is worth highlighting here that the addition of the student level variables reduced

The estimate of the global average was 10,477, revised to 9,617, showing the effect of these variables on the...

student performance. The intra-class correlation remained the same as in the null model (45%).

suggesting that these variables do not explain some of the variation between schools. However, the three variables

The added factors (gender, age, and mother tongue) were significant at the 5% significance level.

($p < 0.05$).

Estimates of the coefficients show that boys perform better.

academic, compared to girls, that is, the fact that the student is male increases the

Performance is 0.14 points lower than women's performance, similar to the result obtained in Laros.

et al. (2010, p. 180), Monteiro (2013, p. 36) and Palermo *et al.* (2014, p. 381). Regarding age, it is possible-

If it is stated that each unit increase in the age of the students causes an increase of 0.21 points in

Academic performance of students, controlled for gender and native language. Regarding attendance.

Regarding how the student speaks Portuguese outside of school, it is observed that the categories "Never ($\bar{y} = \bar{y}1.627$),

Sometimes ($\bar{y} = \bar{y}1.226$) and Often ($\bar{y} = \bar{y}0.645$)", exhibit a negative effect and

significant at the 5% significance level regarding students' academic performance in relation to

The category "Always". Looking at the parameter estimates, it is noticeable that this effect decreases.

as one point increases on the measurement scale of this variable, remembering that the scale

The options used were: 1 – Never; 2 – Sometimes; 3 – Often; and 4 – Always. The model with student variables.

It managed to explain 3.6% of the variation in academic performance (grades) within schools and 3.4% of the...

variation in grades between schools.

Below are the expressions used to calculate the explained variance for each level of analysis:

$$R^2 \text{ Level 1} = \frac{\text{Variance of Model 1} - \text{Variance of Model 2}}{\text{Variance of Model 1}}$$

$$R^2 \text{ Level 2} = \frac{\text{Variance of Model 1} - \text{Variance of Model 2}}{\text{Variance of Model 1}}$$

Once the level 1 variables were added, the next step was to add the level 2 variables.

Level 2, in model 3. The following variables were entered: *the nature of the school, the location of school, type of school, teacher's age, teacher's academic level, years of experience*

The teacher's workload and the teacher/student ratio are used to assess whether these factors explain...

the variability in intercepts (means) between schools, in the academic performance of students. In this

In this stage, all the categorical variables from level 2, mentioned here, were transformed into variables.

dummy (coding 0 for category of no interest and 1 for category of interest). This allows

that the researcher has the freedom to manipulate the variables according to their objective, because

For example, *defining reference categories, exploring interactions, or working only with categories.*

specifics. It is observed that when a variable is not transformed into *dummy variables*, there is no

The possibility of managing your categories, because they are all grouped under the original variable.

The first aspect verified in this model is the model fit, which showed an improvement.

in relation to the previous model, with only student variables (38481.536 < 38777.755). The test

The significance of the model also shows that the improvement from the current model is significant.

(38777.755 - 38481.536 = 296.219). Comparing this value with the critical chi-square value with 20 degrees

of freedom, one can arrive at this conclusion (296.219 > 31.41).

However, none of the school variables included produced any significant effect.

regarding the students' academic performance, or in other words, no variable was statistically significant.

significant at the 5% significance level ($p > 0.05$) with the exception of Type 1 school (school

with more than 1500 students), with a negative effect ($\hat{\gamma} = \hat{\gamma}0.939, p < 0.05$). Thus, it can be stated

that having a larger number of students in school negatively impacts their academic performance, in

approximately one (1) value.

Another aspect that deserves attention in this model is the amount of variance that is explained.

by the model. Observing the results in the Model 3 column in Table 3 (page 8), it can be seen that the

The variance of level 1 and level 2 showed a reduction in their estimates compared to the null model.

The variance estimate for level 1 decreased from 8.373 to 8.090, and the variance estimate for level

2 decreased from 6,894 to 6,149. With this reduction, the portion of the variance explained can be calculated.

at the student level (3.4%) and at the school level (10.8%). With these values, it is noted that the addition of

Contextual variables (high-level variables – school) reduced the amount of variance explained.

Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

observed within schools and increased the amount of variance explained, observed between schools, as was literally expected.

Taking model 2 as a reference, an increase in the variance estimate is observed.

The student level increased from 8,074 in model 2 to 8,090 in the current model (Model 3). This increase in Variance indicates that the model is not explaining performance variability in the best possible way.

among students within schools, suggesting the existence of more variation not explained by the model among students within schools (R2 = -0.2%). However, regarding level 2, the model is able to explain 7.6% of the variation in student performance on the literacy proficiency test. All results

They are in the Model 3 column in Table 3 below.

Table 3. Model without explanatory variables (M1) and with level 1 (M2) and level 2 (M3) variables included.

Parameter	Model 1		Model 2		Model 3	
	Random intercept model without explanatory variables		Model with level 1 variables of random intercepts		Model with level 2 variables of random intercepts	
Level 1 Variables (Student)						
Fixed Effects	EP Estimate	P-value	Estimate	P-value	Estimate	P-value
Intercept	10.484206	0.166	0.000	9.616956	0.291966	0.000
Sex = Male				0.138585	0.067744	0.042
Sex = Female				.	.	.
LM = Never				-1.627491	0.12335	0.000
LM = Sometimes				-1.226489	0.131057	0.000
LM = Often				-0.644958	0.142311	0.000
LM = Always				.	.	.
Age				0.206292	0.022166	0.000
Level 2 Variables (School)						
Type 1 school					-0.939366	0.47058
					0.047	
Random Effects	EP Estimate	P-value	Estimate	P-value	Estimate	P-value
Level 1 variance	8.372602	0.137	0.000	8.073614	0.132371	0.000
Level 2 variance	6.893876	0.639	0.000	6.657812	0.619308	0.000
Variance of the genus						
LM variance						
age variance						
Model Adjustment Measures						
Intraclass Correlation (ICC)	45%		45%		43%	
Deviance (-2LL)	39186,593		38777,755		38481,536	
Parameters	3		8		23	
Deviance 1 - Deviance 2			408,838			
Deviance 2 - Deviance 3					296,219	
Degrees of freedom			5		15	
Chi-square			11,070		24,996	
Variance explained at level 1			3.6%		3.4% (-0.2%)	
Variance explained at level 2			3.4%		10.8% (7.6%)	

Next, all variables that were not... were removed (one at a time).

Significant in Model 3, and it was estimated again in the fourth step with the addition of the effect.

The randomization of the slopes allowed us to test the relationship between the explanatory variables *gender and age*. *Mother tongue* and the dependent variable *academic performance* vary from school to school.

This step goes beyond simply testing the differentiated effect of slopes along the schools.

It also aimed to test whether the presence of the random effect of inclinations would improve the quality of

Adjustment of the model estimated in the previous step (Model 5).



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

Therefore, the results obtained show that the inclusion of the random effect of the slopes

The model fit quality improved, which can be seen in the deviance value, which decreased.

from 38777.755 to 38229.778. The improvement of this model compared to the previous model (Model

5) was significant, since the difference between the deviances of both models (547,977) is greater than

chi-square value with 3 degrees of freedom [$\chi^2(3) = 7.815$]. A problem that was observed

After including the random effect of inclinations, it relates to the significance of the gender variable.

which increased from 5% to 10% ($p = 0.95$). All other variables were significant at the level of

significance of 5% ($p < 0.05$), as can be seen in Table 4 below.

It is observed that the effect of this variable on performance increases with each increase of

a unit on its measurement scale, that is, the more Portuguese the student speaks outside of school,

The higher the student's academic performance, the better. From this result, it can be concluded that: (i) the fact that the student

Never having spoken Portuguese outside of school reduces their performance by 1.57 points compared to

(i) having always spoken; (ii) students who speak Portuguese outside of school sometimes have a reduction in their

performance in 0.95 values compared with students who always speak; (iii) already, for those who

They often speak of this, but the effect reduces their performance to 0.52 points, compared to those who

They always talk.

Corroborating this result, David (2014, p.14) emphasizes that the language and culture of

"Primary socialization" is one of the variables with a notable influence on academic performance.

of the students. Therefore, students from disadvantaged social backgrounds, in which the language of

Socialization and moral values that are contrary to those required by the school would be in better conditions.

unfavorable to responding to the demands of the school system. In addition to this, Freire (2022, p.15)

shows that speaking a different language, or one with some differences, from the one taught in school,

It means having a connection to a particular group whose values and attitudes may not coincide, not

They can be either identity-based or common, along with those of the school culture.

Timbane (2014) emphasizes that the Portuguese language is the basis for students' success in

in the remaining curricular subjects, and those who do not master this language are condemned to failure in the others.

other subjects such as. And according to Gonçalves and Diniz (2004, p. 2) "if students do not master

In this language, the understanding of the content of these subjects is seriously compromised.

hindering their success in school life."

According to Timbane (2009, p. 43), "the child who has Portuguese as their language

A mother enters school after already having traveled a long linguistic path and having developed many...

"communication skills." Therefore, she has an advantage and may even become demotivated by

methodology used with children whose mother tongue is a Bantu language. This student, whose language

Her native language is Portuguese, and she brings with her a background of basic knowledge that allows her to...

to communicate more skillfully than someone who arrives with a different language. This attitude

Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

It can frustrate the student if the appropriate pedagogical measures are not taken.

These results show statistically significant evidence that the government continue with the policy of combating school failure (bilingual education), especially your maintenance in rural areas, where the Portuguese language is seen as a foreign language, and thus, It will be possible to reduce student dropout rates and promote greater inclusion and appreciation of these students. of the local languages (Bantu languages).

Age had a positive effect, increasing students' grades by 0.19.

However, it was found that all the variables considered presented a regression coefficient with random effect, suggesting that the effect of age and mother tongue on the academic performance of Student performance varies between schools, meaning the relationship between these variables and academic performance varies. from school to school ($p < 0.05$).

Table 4. Models with the explanatory variables from levels 1 and 2 included.

Parameter	Model 4			Model 5			Model 6		
	Model with non-significant level 2 variables removed			Model with the Type 1 school variable removed.			Model with random intercepts and slopes removed		
Level 1 Variables (Student)									
Fixed Effects	I estimated	EP	p-value	I estimated	EP	p-value	I estimated	EP	P-value
Intercept	10.043257	0.406559	0.000	9.616956	0.291966	0.000	9.657288	0.318825	0.000
Sex = Male	0.138302	0.067744	0.041	0.138585	0.067744	0.041	0.175352	0.104794	0.095
Sex = Female
LM = Never	-1.627245	0.12334	0.000	-1.627491	0.12335	0.000	-1.567406	0.20993	0.000
LM = Sometimes	-1.226097	0.131049	0.000	-1.226489	0.131057	0.000	-0.946522	0.218338	0.000
LM = Often	-0.645377	0.142306	0.000	-0.644958	0.142311	0.000	-0.517582	0.235793	0.029
LM = Always	.	.	0.000
Age	0.2064	0.022164	0.000	0.206292	0.022166	0.000	0.18591	0.025	0.000
Level 2 Variables (School)									
Type 1 school	-0.566005	0.377814	0.135
Random Effects	I estimated	EP	p-value	I estimated	EP	p-value	I estimated	EP	P-value
Level 1 variance	8.073646	0.132371	0.000	8.073614	0.132371	0.000	6.829602	0.118508	0.000
Level 2 variance	6.597706	0.614076	0.000	6.657812	0.619308	0.000	4.247989	0.705122	0.000
Variance of the genus	0.721262	0.114831	0.000
LM variance	1.744848	0.198092	0.000
age variance	0.025872	0.005732	0.000
Model Adjustment Measures									
Intraclass Correlation (ICC)	45%			45%			38%		
Deviance (-2LL)	38775.52			38777.755			38229.778		
Parameters	9			8			11		
Deviance 3 - Deviance 4	545.74								
Deviance 4 - Deviance 5				-2,235					
Deviance 5 - Deviance 6							547,977		
Degrees of freedom	-14			- 1			3		
Chi-square	NA			NA			7,815		
Variance explained at level 1	3.6% (-0.21%)			3.6% (0.0004%)			18.4% (15.4%)		
Variance explained at level 2	4.3% (-8.2%)			3.4% (-0.9%)			38.4% (36.2%)		

Model 6 further shows that by assuming that inclinations vary from school to school The variances in student level and school level decreased consistently. Therefore, making use of the coefficient of determination R^2 , adapted to the multilevel model, it is observed that After including the random effect of the slopes, the model registered a tremendous improvement in... its explanatory power. Model 6 explained 38.4% of the variability between schools and 18.4% of the variability within schools. Considering Model 5 as a point of reference, this model It explains 36.2% of the variance at level 2 and 15.4% of the variance at level 1.

**FINAL CONSIDERATIONS**

Up to this point, research shows that the use of the multilevel analysis approach in Mozambique is scarce. Therefore, it is believed that this study has sparked interest among students and Researchers from various fields of knowledge recognize the need to align themselves with this new approach. analysis in such a way that they will then look at problems from a more complex perspective. analysis.

The study showed that among the variables considered, age and mother tongue are the variables that influence students' academic performance, with a positive effect.

Since student performance is strongly influenced by their native language (medium of (socialization), alongside bilingual education, community programs should be implemented. from teaching reading, because, according to several authors, this is where the biggest problem lies. One can also consider other rural socio-economic development programs, because in Most of the time, a student's academic performance is associated with their level of development. socio-economic status of their social circle, according to Timbane (2014).

For children who are not yet attending any classes, parents who are able to They should invest in their integration into preschool education as early as possible. Studies such as those by UNICEF (2021, p. 9) show that the frequency of pre-primary education increases the probabilities of children can "perform well," regardless of their age or gender. This strategy could to help control the significant effect of age, which in this study was positive.

References

- AITKIN, M.; ANDERSON, D.; HINDE, J. *Statistical modeling in school effectiveness*. Journal of the Royal Statistical Society: Series A (General), London, v. 144, p. 1–43, 1981.
- AITKIN, M.; LONGFORD, NT *Statistical modeling issues in school effectiveness studies*. Journal of the Royal Statistical Society: Series A (General), London, v. 149, p. 1–43, 1986.
- AMÉRICO, BL; LACRUZ, AJ. *Context and school performance: analysis of grades in the Prova Brasil of schools in Espírito Santo using multiple linear regression*. Revista de Administração Pública, Rio de Janeiro, v. 51, n. 5, p. 854–878, 2017.
- BENNET, N. *Teaching styles and pupil progress*. London: Open Books, 1976.
- COLEMAN, JS *Equality of educational opportunity*. Washington, DC: Office of Education, 1966.
- COSTA, GDF. *Relationships between motivational orientations and the academic performance of 7th grade elementary school students in mathematics, in solving first-degree equations*. Dissertation (Master's in Education) – State University of Campinas, Campinas, 2005.
- DAVID, MB *Bilingual education in Mozambique: a sociological analysis of the influence of context and students' social origins on their school performance at EPC Lussa, Gurúe District*,



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

Zambezia Province. Monograph (Bachelor's Degree in Sociology) – Eduardo Mondlane University, Maputo, 2017.

FARIA, EM; GUIMARÃES, RRM *Excellence with equity: school factors for educational success in unfavorable circumstances.* Studies in Educational Evaluation, São Paulo, v. 26, n. 61, p. 3–25, 2015.

FERRÃO, ME. *Introduction to multilevel regression models in education.* Campinas: Komedi, 2003.

FERRÃO, ME; FERNANDES, C. *Multilevel model: an application to educational assessment data.* Estudos em Avaliação Educacional, São Paulo, n. 22, p. 135–153, 2000.

FIELD, A. *Discovering statistics using IBM SPSS statistics.* 4th ed. Thousand Oaks, CA: Sage, 2013.

FONTES, M. *School performance.* Knoww.net, 2017.

FREIRE, CM. *The mastery of the Portuguese language as an influencer of the school results of students in the 3rd cycle with origin or ancestry in the PALOP and Brazil.* Dissertation (Master's in Education Studies) – Universidade Nova de Lisboa, Lisbon, 2022.

GOMES, MM. *Factors that facilitate and hinder learning.* Public Education Journal, Rio de Janeiro, 2018.

HAYDT, RC. *Evaluation of the teaching-learning process.* 6th ed. São Paulo: Ática, 1997.

HECK, RH; THOMAS, SL; TABATA, LN *Multilevel and longitudinal modeling with IBM SPSS.* 2nd ed. New York: Routledge, 2014.

HOX, JJ *Applied multilevel analysis: techniques and applications.* Amsterdam: TT Publikaties, 1995.

HOX, JJ *Multilevel analysis: techniques and applications.* Mahwah, NJ: Lawrence Erlbaum Associates, 2002.

JESUS, GR DE; LAROS, JA *School effectiveness: multilevel regression with evaluation data.* Psychological Assessment Journal, Porto Alegre, v. 3, n. 2, p. 93–106, 2004.

KREFT, I.; DE LEEUW, J. *Introducing multilevel modeling.* London: Sage, 1998.

LAROS, JA; MARCIANO, JLP. *Multilevel analysis applied to NELS:88 data.* Estudos em Avaliação Educacional, São Paulo, v. 19, n. 40, p. 3–21, 2008.

LAROS, JA; MARCIANO, JLP; ANDRADE, JM DE. *Factors affecting performance on the SAEB mathematics test: a multilevel study.* Revista Avaliação Psicológica, Porto Alegre, v. 9, n. 2, p. 173–186, 2010.

LICUCO, PM *Application of multivariate techniques for the identification of determining factors in pedagogical achievement.* Monograph (Bachelor's Degree in Mathematics and Informatics) – Eduardo Mondlane University, Maputo, 2011.

LUZ, LS *The determinants of school performance: educational stratification and the value effect*



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

Added. In: NATIONAL MEETING OF POPULATION STUDIES, 15th, 2006, Caxambu.

Proceedings. Caxambu: ABEP, 2006.

MACEDO, GA. *Factors associated with the academic performance of 5th grade students: a value-added approach*. In: NATIONAL MEETING OF POPULATION STUDIES, 14th, 2004, Caxambu. Proceedings. Caxambu: ABEP, 2004.

MAROCO, J. *Statistical analysis with SPSS Statistics*. 5th ed. Lisbon: Report Number, 2011.

MENEZES-FILHO, NA. *The determinants of school performance in Brazil*. São Paulo: Instituto Futuro Brasil, 2007.

Ministry of Education and Human Development. *National Evaluation: Management Questionnaire*. 2nd ed. Maputo: MINEDH, 2016.

Ministry of Education and Human Development. *Strategic Education Plan 2020–2029: Education Sector Performance Report 2021*. Maputo: MINEDH, 2022.

MONTEIRO, MASPD *Multilevel regression models: an application in education*.

Dissertation (Master's in Statistics) – University of Minho, Braga, 2013.

NATIS, L. *Hierarchical linear models: construction and interpretation of knowledge scales*.

Studies in Educational Evaluation, São Paulo, n. 23, p. 3–29, 2001.

PALERMO, GA; SILVA, DBN; NOVELLINO, MSF. *Factors associated with school performance*. Brazilian Journal of Population Studies, Rio de Janeiro, v. 31, n. 2, p. 367–394, 2014.

RAUDENBUSH, SW; BRYK, AS *Hierarchical linear models: applications and data analysis methods*. 2nd ed. Thousand Oaks, CA: Sage, 2002.

SNYDERS, TAB; BOSKER, RJ *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. 2nd ed. Thousand Oaks, CA: Sage, 2012.

SOARES, JF. *The effect of school on the cognitive performance of its students*. Ibero-American Journal on Quality, Effectiveness and Change in Education, v. 2, n. 2, p. 83–104, 2004.

TIMBANE, AA. *The problem of teaching the Portuguese language in the 1st grade in an urban sociolinguistic context*. Dissertation (Master's in Education) – Eduardo Mondlane University, Maputo, 2009.

TIMBANE, AA. *The teaching of the Portuguese language in Mozambique and the problems of primary school teacher training*. Artíficos Magazine, Ilhéus, v. 4, n. 7, 2014.

TWISK, JWR *Applied multilevel analysis*. New York: Cambridge University Press, 2006.

UNICEF. *School attendance and performance of primary school children in Mozambique*. Maputo: UNICEF, 2020.

UNITED STATES AGENCY FOR INTERNATIONAL DEVELOPMENT. *Improving Learning Outcomes in Primary Education (SABER)*. Maputo: USAID, 2021.



Year VI, v.1 2026 | Submission: 09/01/2026 | Accepted: 11/01/2026 | Publication: 13/01/2026

VERNIER, LDS; BAGOLIN, IP; JACINTO, PA. *Factors influencing school performance in the state of Rio Grande do Sul.*
Revista de Economia e Sociologia Rural, Porto Alegre, v. 53, n. 1, p. 63–78, 2015.

Financing

This work was carried out with the financial support of the Government of Mozambique and ICIPE – World Bank, REF: RSIF/RA/022 for the PASET Regional Fund for Scholarships and Innovation. The opinions expressed here do not necessarily reflect the official opinion of the donors.