ISSN: 2317-0840

Association between scientific initiation and success in postgraduate studies

Douglas Oliveira Gomes[1] , Helgem de Souza Ribeiro Martins[1†], Josino José Barbosa[1] , Gabriel Lima de Souza[1] , Anderson Ribeiro Duarte[1] .

[1]*Universidade Federal de Ouro Preto, Instituto de Ciências Exatas e Biológicas, Departamento de Estatística; Ouro Preto-MG, Brasil.*

Abstract: *This study investigates the relationship between participation in Scientific Initiation (SI) and admission and success rates in graduate programs at the Federal University of Viçosa (UFV). The research was motivated by the need to understand the factors that influence access and academic continuity, especially in the Brazilian context, where graduate school dropout rates pose a challenge for both students and institutions. Using descriptive statistics and logistic regression techniques, data were obtained through the Federal University of Viçosa's open data portal, and specific data were analyzed using R software. Variables such as age, gender, state of origin, and academic standing were examined, enabling the identification of patterns and trends in student trajectories. It was possible to draw an average profile of UFV students, both at the undergraduate and graduate levels, and to assess the relationship between scientific initiation and graduate studies. The results indicate that participation in scientific initiation has a positive association with admission to graduate studies. Students who participated in scientific initiation during their undergraduate years had a significantly higher rate of academic continuation compared to those who did not participate in this program. The results of this research underscore the importance of institutional policies that encourage scientific initiation, promoting greater academic engagement and contributing to a reduction in dropout rates in graduate programs. The study also contributes to the knowledge base on the relationship between scientific initiation and retention in graduate programs.*

Keywords: *Postgraduate studies; Scientific initiation; Data science; Logistic regression; School dropout.*

Introduction

Scientific Initiation (SI) is a training tool for undergraduate students, aiming to introduce them to the world of scientific research. According to Bastos (2010), scientific initiation is a theoretical and methodological support tool that enables the implementation of a project that directly contributes to the student's professional development.

The practice of initiating undergraduate students into scientific projects dates back to the 19th century. German universities incorporated scientific research into the undergraduate training process to develop them into more capable professionals and teachers, endowed with scientific thinking. Based on this model, the current practice, in which students are mentored by a supervisor versed in scientific methodology and a specific field of study, was refined (Lordelo; Argôlo, 2015).

During the scientific initiation project, students become familiar with the theoretical aspects of scientific methodology, including the general scientific process, writing scientific papers, and conducting research. At the same time, they are introduced to research methodologies, which fosters their development as scientists and future professionals, as this phase exposes them to the latest developments in their respective fields of study.

[†]Autor correspondente: helgem@ufop.edu.br
Received: 25/08/2025. Revised: 09/10/2025. Accepted:10/10/2025.

Beyond the aspect related to scientific training, scientific initiation in Brazil provides financial resources to students who participate in this modality. According to Massi (2010), funding for scientific research at the undergraduate level began to be offered in the 1950s, with the creation of the National Council for Scientific and Technological Development (CNPQ). The council awarded annual scholarships to promote undergraduate research.

Beyond the aspect related to scientific training, scientific initiation in Brazil provides financial resources to students who participate in this modality. According to Massi (2010), funding for scientific research at the undergraduate level began to be offered in the 1950s, with the creation of the National Council for Scientific and Technological Development (CNPQ). The council awarded annual scholarships to promote undergraduate research.

Currently, funding for undergraduate scientific research comes from a variety of sources. In addition to the CNPq (National Council for Scientific and Technological Development), other organizations promote such activities, such as research support foundations (FAPs), non-governmental organizations, and non-profit institutes, companies, and higher education institutions themselves, particularly IFES (Federal Higher Education Institutions), which have internal funding for scientific initiation scholarships. It is common for students participating in scientific initiation programs to develop such a deep interest in science that they decide to pursue their education at the postgraduate level. This pattern is considered a natural path within academia. Due to this natural progression in the development of a researcher, several studies have examined this relationship.

The study by Silva (2019) addresses the impact of students' participation in scientific initiation programs on their postgraduate academic careers. The work by Dantas (2019) addresses a similar topic, this time within the Federal University of Rio Grande do Norte (UFRN), using a census approach.

The study by Lordelo (2015) analyzed the impact of scientific initiation programs (SI) on postgraduate studies at a federal institution in Brazil, comparing 895 master's students (242 SI participants and 653 non-SI participants). The analysis, based on data from the Lattes Platform, suggests that SI facilitates the transition to graduate school and enhances research engagement.

The article by Pelinski (2021) examines the impact of scientific initiation (SI) programs on graduate admissions at UEPG. The study, which employed descriptive statistics and Propensity Score Matching (PSM) using the Logit model, demonstrated that SI has a significant impact on admission rates, particularly for students in the exact, biological, and applied social sciences.

The study conducted by Lima (2016) investigated the differences between students who graduated from the UFRGS graduate program in 2012, focusing on the impact of their experience in Scientific Initiation (SI) during their undergraduate studies. The analysis, based on descriptive statistics, compared two groups: those with SI experience (SI Group) and those without (NIC Group). The research revealed that the IC Group graduated younger and in less time than the NIC Group, in addition to presenting greater scientific production in several areas.

The Federal University of Viçosa offers several scientific initiation programs, some of which are funded by agencies such as CNPq and FAPEMIG, in addition to a program for voluntary participation in scientific initiation. In addition to scientific initiation programs, UFV boasts an ecosystem of graduate programs with a wide variety of programs in various fields of knowledge, particularly in the stricto sensu modality, offering master's and doctoral programs.

Thus, there is significant interaction between scientific initiation and graduate programs at UFOP. Many scientific initiation students even form part of the research project teams of master's and doctoral students.

Therefore, to assess the level of interaction between these activities and to measure the impact of scientific initiation activities on the admission and performance of graduate students at this institution, there is institutional interest in this analysis.

The primary objective of this study was to assess, through quantitative analysis, the participation in scientific initiation programs and their impact on the admission and performance of UFV graduate students.

The specific objectives of this study were as follows:

- To collect data related to scientific initiation and graduate programs at UFV;
- To verify the number of UFV scientific initiation students who choose to continue their studies in graduate programs at UFV;
- To evaluate the impact of participation in scientific initiation programs on the admission and performance of UFV graduate students.
- To develop a methodology for ongoing monitoring of scientific initiation students in UFV graduate programs.

Materials and Methodological Tools

This section will cover details on database access, data preparation, and the methodological tools applied to the problem.

Databases

This study utilizes two databases that contain information related to undergraduate and graduate students from the Federal University of Viçosa. The data were obtained directly from the institutional open data portal and are available at the following link: ⟨https://dados.ufv.br/⟩.

The first database under study comprises seven variables and 20,825 observations of participation in Scientific Initiation (SI) programs at UFV, spanning the period from 1990 to 2025. It includes variables such as student identifier, course, project title, advisor, scholarship type, and start and end dates.

The second database analyzed contains 23 variables and 96,091 observations on students at the Federal University of Viçosa (UFV), spanning the period from 2015 to 2024. The recorded variables include demographic data, such as gender, year of birth, race, place of birth, state of origin, and country, as well as information on the method of admission to the university (entrance exam, diploma holder, among others) and whether the student belongs to a quota group. Academic data are also presented, such as undergraduate program, grouped level, program area, shift, student status (completion, dropout, withdrawal), and campus.

The strategy adopted consisted of unifying two distinct databases by using a unique identifier present in both to track which students were in both sources. This process allowed identifying those who participated in Scientific Initiation (SI) during their undergraduate studies and subsequently attended graduate school. Because one database had a significantly larger volume of records than the other, temporal planning was performed, considering only the years 2010 to 2024 and storing them in the ”AdmissionYear” variable. This filtering was crucial in making the analysis more accurate and manageable. After this step, the databases were combined with relevant information from both sources, which facilitated the analysis.

After describing the final database, several transformations and analyses were applied to facilitate data interpretation and understanding. First, the age variable was created, obtained

by subtracting the year of birth from the current year (2024). Then, a variable age range was generated to segment students into specific categories: 0-18, 19-21, 22-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-60, and 60+.

Additionally, the SI variable was created using the unique identifier present in both databases to verify the student's participation in the Scientific Initiation program. This variable was assigned the value "Yes" for students present in both databases and "No" for all others. Similarly, the postgraduate variable was generated to indicate whether the student entered graduate school, with a value of "Yes" for those identified in the postgraduate database and "No" for those who did not. To facilitate analysis, the undergraduate variable was included, which assumes the value "Yes" for all students, as the database was constructed based on undergraduate data.

To improve the analysis by state of origin (UF), states with low representation were grouped into the "Other states" category, while records without information were classified as "NI" (Not Reported). Finally, the nationality variable was created, differentiating Brazilian and international students based on their country of origin.

Methodological tools

This work employed descriptive analysis and logistic regression as the primary tools to address the questions raised in the objectives section.

Descriptive statistics is one of the most important branches of statistics. Its objective is to organize and summarize the data under analysis to facilitate their interpretation and indicate possible subsequent analyses. Such analyses are based on numerical and graphical methods (Montgomery, 2017).

The analysis of the relationship between the variables under study will use a logistic regression model. Logistic regression is the primary inferential statistical analysis technique used in analyzing dichotomous data. Its primary purpose is to model the probability associated with success/failure events, based on explanatory variables Agresti (2011). The logistic regression model is represented by:

$$\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p \tag{1}$$

where:

- $(\beta_0, \beta_1, \ldots, \beta_p)$ - Vector of parameters
- $(1, x_1, x_2, \cdots, x_p)$ – Vector of explanatory variables
- $\pi(x)$ - Probability of success associated with the model

The basis for interpreting this model is the exponential of the coefficients, which represents the odds ratio of success to failure. Significant variables were selected using the stepwise selection method, with a significance level of 1

The first stage of analysis consisted of a descriptive analysis to examine the distribution of students based on demographic and academic variables, such as gender, age, state of origin (UF), and nationality, with the use of bar charts and cross tabs, to visualize these distributions and identify possible patterns in student composition. The same techniques can be used to assess the differences between undergraduate and graduate student groups and to investigate the influence of participation in Scientific Initiation (SI) on academic progression, as well as the impact of this variable on retention and course completion.

Subsequently, the descriptive evidence was analyzed from an inferential perspective, using logistic regression techniques, to quantify the association between participation in scientific initiation programs, admission, and student performance in postgraduate programs.

Results and Insights

The objectives of this study were to investigate the impact of participation in Scientific Initiation (SI) on students' academic continuity, analyzing its influence on admission to and retention in graduate school. Furthermore, the study aimed to provide a comprehensive overview of the student body's profile over the years, enabling a better understanding of the factors that may impact academic continuity.

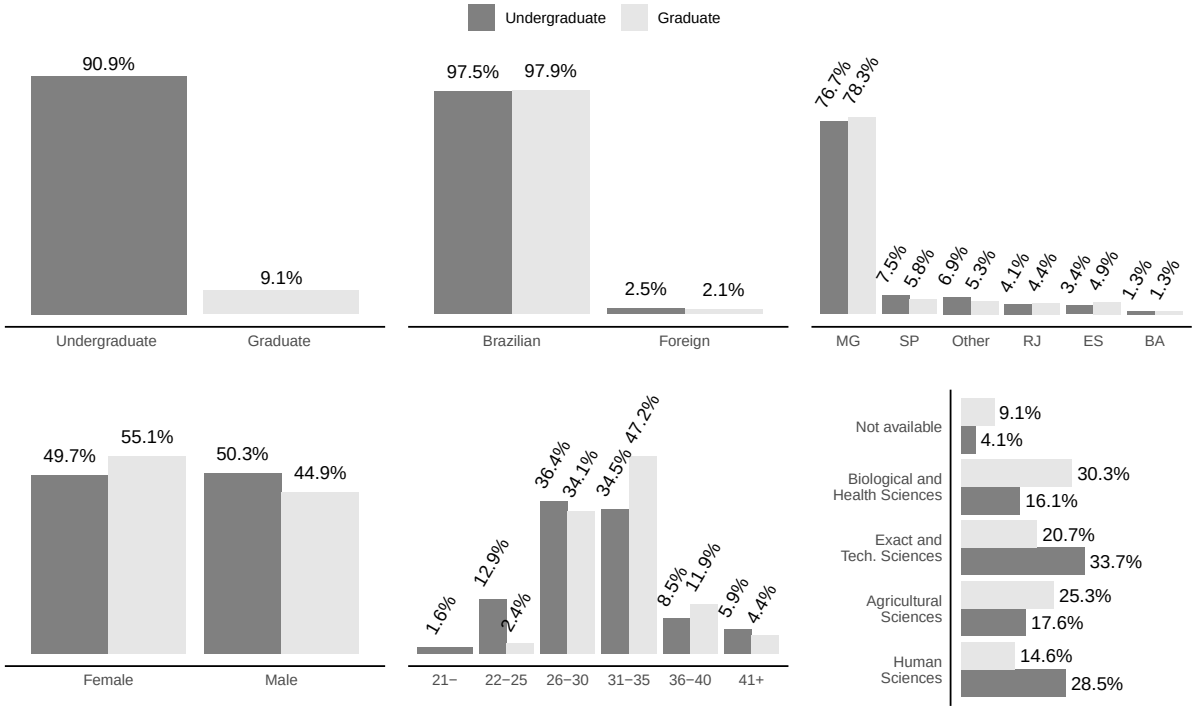
To achieve these objectives, descriptive analyses were conducted, allowing for the characterization of students in terms of gender, age, state of origin (UF), and nationality. Furthermore, comparative analyses were conducted, allowing for the examination of the relationship between SI participation and academic progression, identifying patterns of admission, retention, and dropout in graduate school. The unification of the databases was essential to facilitate this investigation, enabling the precise tracking of students throughout their academic trajectory.

The following sections present the results of these analyses and discuss their implications within the institutional and academic context, providing insights that can contribute to the development of more effective educational policies and the strengthening of scientific training in Brazil.

Sample Characterization

This section describes the main characteristics of the sample, such as the distribution of students by age, gender, state of origin, and nationality. This analysis allows us to understand the student profile at UFV and informs research on the relationship between participation in Scientific Initiation and admission to graduate programs. The results will be presented by stratum, considering the course level (undergraduate or graduate). This allows for an understanding of both the overall profile and the profile by level. The results are presented in the Figure 1.

Figure 1: Sample characterization by course level



Source: from the authors (2025).

The vast majority of the institution's students are undergraduates, accounting for approximately 91% of the total. Regarding nationality, the absolute majority of students are Brazilian, with very similar proportions at both levels. Regarding state of origin, it is observed that over 75% of students come from the state of Minas Gerais (MG), where the institution is located. The second state with the most students is São Paulo (SP), a neighboring state to Minas Gerais. The other states represented (Rio de Janeiro (RJ) and Espírito Santo (ES)) have a significant proportion of students. It is worth noting that these states are located in the same geographic region as Minas Gerais, which facilitates access to the institution. The combined number of students from other Brazilian states accounts for approximately 7% of the total undergraduate students and 5% of the graduate students.

Regarding gender, there is a balance between male and female students in undergraduate programs. However, in graduate programs, the female presence is more prominent, with approximately 10% more female students compared to male students, indicating a greater tendency for women to continue their studies at the graduate level. Regarding age, a typical and expected pattern is observed: younger students are enrolled in undergraduate programs. As age increases, the relationship reverses, and older students predominate in graduate programs. A particularly interesting characteristic is that students aged 41 or older (41+) are included. From this age onward, the relationship reverses again, with a higher proportion of undergraduate students than graduate students, which may indicate a trend toward returning to studies at this age.

The proportion of students by field of study also shows interesting patterns. The fields of Exact Sciences, Technology, and Engineering, along with the Humanities, have a higher proportion of undergraduate students compared to graduate students. In the fields of Biological and health sciences, as well as agricultural sciences, the relationship is reversed, which may indicate that these areas are where institutional research is more advanced.

Relationship between Scientific Initiation and Postgraduate Studies

This section presents the results obtained from analyzing the relationship between participation in Scientific Initiation (SI) and admission to graduate programs (GP). The objective is to identify possible patterns and trends that indicate the influence of SI on the continuation of graduate studies, as well as to assess whether students who participated in this program and entered graduate programs were more likely to complete their studies, compared to those who did not participate.

Table 1: Table of Variables Related to Postgraduate Admission and Completion.

Variable	**Admission to GP**		**Complete GP**	
	No	Yes	No	Yes
	N = 44,306	N = 9,007	N = 3,022	N = 5,985
Scientific Initiation				
No	40,108 (88%)	5,214 (12%)	1,973 (38%)	3,241 (62%)
Yes	4,198 (53%)	3,793 (47%)	1,049 (28%)	2,744 (72%)

[1] n (%)

Source: From the authors (2025).

Table 1 presents, descriptively, the relationship between participation in graduate programs, the eventual entry of students into graduate programs, and the completion of those who do enter these programs. Regarding admission to graduate programs, the relationship is quite evident. Among students who did not participate in research initiation programs, only 12% of those observed during the period in question entered graduate school. Among those who participated in this preparatory stage, the percentage rises to 47%, indicating that research initiation constitutes a strong incentive for continuing studies at the graduate level.

When considering students who entered graduate programs, it is clear that the proportion of students who completed their studies is higher among those participating in research initiation programs, with 62% of graduates from programs not participating in research initiation programs compared to 72% of graduates among those who did. This evidence indicates that initial training in scientific research can not only favor admission but also contribute to the successful development of studies.

To conclude the results section, we will present the results of two logistic regression models to investigate the relationship between scientific initiation and admission to, as well as completion of, graduate programs. The results are presented in the tables below.

The first logistic regression model, whose results are available in Table 2, examines the association between participation in research initiation programs and admission to graduate programs. The model confirms the sample evidence, indicating that a student with research initiation experience is approximately seven times more likely to enter graduate school than a student without such experience.

Table 2: Associations with Postgraduate Enrollment.

	Admission to GP		
Variable	**Odds Ratio**	**Confidence Interval (95%)**	**p-value**
Scientific Initiation			
No	–	–	–
Yes	6.95	6.59, 7.33	<0.001

Source: From the authors (2025).

Regarding the association between students' prior experience in science and their graduation rate after entering graduate school, we found a significant relationship between the variables. The results are available in Table 3. Students with research initiation experience who were able to access graduate school had approximately a 60% higher chance of graduation compared to those who were not research initiation students.

Table 3: Associations with Postgraduate conclusion.

	Conclusion of GP		
Variable	**Odds Ratio**	**Confidence Interval (95%)**	**p-value**
Scientific Initiation			
No	–	–	–
Yes	1.59	1.45, 1.74	<0.001

Source: From the authors (2025).

The results consistently demonstrate that, among students at the Federal University of Viçosa, there is a clear association between participation in scientific initiation programs and outcomes related to graduate studies. The first of these concerns access to graduate programs.

Among students with scientific initiation experience, significantly higher access to graduate programs is observed. These findings are consistent with the results reported by Lordelo (2015), which indicate that scientific initiation accelerates students' transition from undergraduate to graduate studies. The work by Pelinski (2021) also provides a similar conclusion regarding the chances of access.

The other relationship outlined by this study indicates that students with initial scientific research experience at the institution have a higher completion rate compared to those without such experience. The results obtained are consistent with the findings of the study by Lima (2016), which conclude that students with scientific initiation experience generally obtain their postgraduate degrees in a shorter time than others, as well as at a younger age. The study by Lordelo (2015) also presents results in this regard, as it indicates an increase in these students' engagement in research.

In this sense, the results presented by this study, in addition to converging with the results of previous studies, provide a robust analysis of data from a large educational institution, which generates new evidence of the importance of scientific initiation policy both in recruiting students for scientific initiation and in achieving lower dropout rates in postgraduate programs.

The primary limitations of this study stem from its local nature. Due to the difficulty in accessing undergraduate and graduate data, the study relies on access to open data platforms or the provision of data by institutions. However, this study has great potential for generalization.

Final Remarks

This study provides exciting results from a research perspective on metrics that link scientific initiation and graduate studies. Initially, a descriptive analysis allowed for the characterization of undergraduate and graduate students at the Federal University of Viçosa between 2015 and 2024. Furthermore, the work contributes to the expansion of research on this topic.

The study demonstrated the importance of scientific initiation in students' academic careers and provided evidence that those who participated in scientific initiation programs are more likely to enter graduate school. Furthermore, it provides evidence that prior experience in scientific research can increase a student's likelihood of completing their graduate studies.

The results reinforce the need for institutional policies that encourage undergraduate student participation in research initiation programs, which can reduce dropout rates and strengthen academic training at both the undergraduate and graduate levels.

For future studies, we suggest applying the methodology presented to data from other higher education institutions, as well as utilizing new methodologies to predict graduate success, such as those based on machine learning techniques.

Conflicts of interest and the use of artificial intelligence

The authors declare that there is no conflict of interest in conducting this research, and the authors also confirm that no AI resources were used.

Acknowledgements

The authors would like to thank the Universidade Federal de Ouro Preto for partially supporting the development of this study.

References

AGRESTI A.; KATERI, M. Categorical data analysis. In: **International encyclopedia of statistical science**. [S.l.]: **Springer**, 2011. p. 206–208.

BASTOS F.; MARTINS, F. A. M. T. M. L. C. S. A importância da iniciação científica para os alunos de graduação em biomedicina. **Revista Eletrônica Novo Enfoque**, v. 11, n. 11, p. 61–66, 2010.

DANTAS, E. H. F. Formação de recursos humanos para pesquisa: avaliação do impacto do **programa institucional de bolsas de iniciação científica na pós-graduação da ufrn**. 2019.

LIMA, L. G. A. **A influência da iniciação científica sobre a pós-graduação: um estudo de caso sobre tempo, idade de titulação e produção científica**. 2016.

LORDELO, J. A. C.; ARGôLO, R. F. Influências da iniciação científica na pós-graduação. **Estudos em Avaliação Educacional**, v. 26, n. 61, p. 168–191, 2015.

MASSI L.; QUEIROZ, S. L. Estudos sobre iniciação científica no brasil: uma revisão. **Cadernos de Pesquisa**, v. 40, n. 139, p. 173–197, 2010.

MONTGOMERY, D. C. **Design and analysis of experiments**. Hoboken: **John Wiley & Sons**, 2017.

PELINSKI J. D.; KAMPA, A. Efeito da iniciação científica no rendimento acadêmico e na inserção em programas de pós-graduação "stricto sensu": uma análise com propensity score matching. **Universidade Estadual de Ponta Grossa**, 2021.

SILVA P. S. B.; BIN, A. da. O impacto da iniciação científica na pós-graduação. **Seminários do LEG**, n. 8, p. 93–99, 2019.