# IMPACT EVALUATION IN EDUCATION

## The successful experience of the Jovem de Futuro program in partnership with the government

Ricardo Henriques
Mirela de Carvalho
Ricardo Paes de Barros

INSTITUTO UNIBANCO

# IMPACT EVALUATION IN EDUCATION

## The successful experience of the Jovem de Futuro program in partnership with the government

**Ricardo Henriques**
**Mirela de Carvalho**
**Ricardo Paes de Barros**

INSTITUTO UNIBANCO

**PRODUCTION**
Instituto Unibanco

**ADMINISTRATIVE COUNCIL**
**Chairperson**
Pedro Moreira Salles

**Vice-Chairperson**
Pedro Sampaio Malan

**Board Members**
Antonio Jacinto Matias
Claudia Costin
Cláudio de Moura Castro
Cláudio Luiz da Silva Haddad
Marcelo Luis Orticelli
Marcos de Barros Lisboa
Ricardo Paes de Barros
Rodolfo Villela Marino

**Board of Directors**
Cláudio José Coutinho Arromatte
Jânio Francisco Ferrugem Gomes
Leila Cristiane Barboza Braga de Melo
Marcelo Luis Orticelli
Paulo Sergio Miron
Ricardo Urquijo Lazcano

**TECHNICAL STAFF**
**Executive Superintendent**
Ricardo Henriques

**Managers**
Cesar Nunes
Maria Julia Azevedo
Mirela de Carvalho
Tiago Borba
Vinicius da Silva Scarpi

**DEVELOPMENT OF TEXTS**
**Authors**
Ricardo Henriques
Ricardo Paes de Barros
Mirela de Carvalho

**Production of content**
Beatriz Silva Garcia
Fabiana Silva Bento
Gabriel Guimaraes Leite
Laura Muller Machado
Mariana Bittar
Paula Penko Teixeira
Samuel Franco

**Content editing**
Antônio Gois

**EDITORIAL PRODUCTION**
**Coordination**
José Jacinto Amaral

**Revision**
Demetrio Pires Weber Candiota da Rosa
Érico Coelho de Melo

**Graphic Design and Layout**
Fernanda Aoki

**Illustrations**
Mario Kanno

Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)

H519a

Henriques, Ricardo.
Avaliação de impacto em educação: a experiência exitosa do programa Jovem de Futuro em parceria com o poder público / Ricardo Henriques, Mirela de Carvalho, Ricardo Paes de Barros. – São Paulo, SP: Instituto Unibanco, 2020.
56 p. : il. color. ; 16,5 x 24 cm

Inclui bibliografia
ISBN 978-65-86069-53-2

1. Avaliação educacional. 2. Educação e Estado. 3. Programa Jovem de Futuro. I. Carvalho, Mirela de. II. Barros, Ricardo Paes de. III. Título.
CDD 378.81

# TABLE OF CONTENTS

# INTRODUCTION

Effective school leadership is fundamental to student achievement and success. Principals act as critical agents of change within schools, leading faculty, staff, and students to improve teaching and learning. Indeed, an international body of empirical evidence confirms that school leaders substantially impact student achievement. Amongst school-related factors, principals are second only to teachers in terms of their impact on student learning outcomes.[1]

Over the last several decades, researchers have sought to understand the specific leadership practices that influence student achievement the most. In a widely referenced academic study, New Zealander Viviane Robinson and her colleagues identified a strong association between school leaders involved in promoting and participating in teacher professional development and better student outcomes.[2] The U.S. Department of Education's Institute of Education Sciences (IES), the country's leading source of evidence-based information on education programs and policies, investigated the relationship between leadership practices and school improvement, providing further evidence for the importance of school leaders. A study conducted as part of IES' *What Works Clearinghouse*[3] recommended a specific set of actions and practices to drastically improve low-performing schools: (1) strong leadership communicating the need for dramatic change; (2) a consistent focus on improving instruction; (3) visibility for early improvements and quick wins; and (4) school leaders focused on building a staff fully committed to the school's improvement goals.
Globally, policymakers are placing a greater emphasis on school

1 See Leithwood et al., 2004.

2 See Robinson, Lloyd and Rowe, 2008.

3 See Herman et al., 2008.

leaders as catalysts for educational improvement. Scientific evidence has informed the development of the education leadership standards in the United States (Council of Chief State School Officers, 2008), the work of the National College of School Leadership and Children's Services in the United Kingdom (Leithwood; Day; Sammons; Harris; Hopkins, 2006), and the development of leadership frameworks in Australia and New Zealand (New Zealand Ministry of Education, 2008). (Robinson, 2011).[4] Given their influence on students, school leaders are also being held more accountable for achieving meaningful so results. According to Helene Ärlestig, Christopher Day, and Olaf Johansen, authors of *A Decade of Research on School Principals,*[5] an analysis of two dozen countries shows that "school principals in all countries are subject to more and more public scrutiny and held to be more closely accountable by governments for the academic attainment and equity of learning opportunities for all their students."

Despite the substantial body of evidence linking school leadership and improved student outcomes, public policy in Brazil and Latin America has yet to fully address how to improve the skills and capacities of principals. In a study published in 2016, José Weinstein and Macarena Hernández found that school systems across Latin America have attempted to address school leadership, but that these policies are "still at an early stage, with several problems, contradictions, and lack of internal-external coherence."[6]

School leaders matter. High-performing education systems throughout the world focus on developing and supporting principals to achieve results. In Ontario, Canada, for example, policymakers sought to improve school leadership capacity and practice in order to increase student outcomes. In the words of Michael Fullan, the state's former education advisor: "If we see an increase in the best practices and our case studies show that the practices work and student achievement is increasing over the long haul, time and again, then we can be confident about the difference that leadership practice is making. Leadership has made the difference in Ontario – it's undeniable – we can trace it, feel it, see it."[7] The promise of school leadership as a means of improving student achievement motivates the work of Jovem de Futuro.

4 Robinson, 2011.

5 Ärlestig; Day; Johansen et al., 2016.

6 See Weinstein and Hernández, 2016.

7 Ontario Leadership Congress, April 2012. Available at: <http://www.edu.gov.on.ca/eng/policyfunding/leadership/OLSQuickFacts.pdf>. Accessed on: May 8, 2020.

## ABOUT JOVEM DE FUTURO

Jovem de Futuro is a public-private partnership focused on improving leadership and management practices in schools, regional education offices, and state departments of education. The ultimate goal of the program is to develop school leaders and increase the high school graduation rate and ensure that young people are equipped with a sufficient education as they enter adulthood.

The program focuses on high school because of the challenges that exist in Brazilian secondary education. In 2005, only 5% of public school students in their third year of high school were performing on grade-level in mathematics,[8] only seven in every ten public school students passed their end-of year exams, and nearly one in five students dropped out of school.[9] As a result, education leaders and civil society sought to address inequities in education. In 2007, the Brazilian Ministry of Education implemented the Basic Education Development Index (IDEB) to monitor student achievement and student progression towards high school graduation. In the last two decades, Brazil has increased its capacity to collect data on school performance and student outcomes. Access to these data drives policy decisions and education programs across the country.

Jovem de Futuro seeks to improve student outcomes by focusing on school leaders. A large body of academic literature from across the world confirms the importance of effective school leadership on student outcomes. Research shows that principals play an important role--nearly as important as teachers--in improving student outcomes,[10] and evidence suggests that school leaders can improve student learning by focusing on teacher development and pedagogical improvement. Principals in Brazil must reimagine their role as instructional leaders rather than bureaucratic managers.

## IMPACT ASSESSMENT SINCE THE BEGINNING

Jovem de Futuro has been committed to evaluating its impact using rigorous research methods since the inception of the program. Because of this commitment, researchers evaluated the program using a randomized experimental approach. This method randomly assigns participants to two groups (one that participates in the program and another that does not), ensuring that the only expected difference between individuals in the two groups is their participation--or lack

8 According to the scale created in 2006 by the technical commission of the '*Todos Pela Educação*' ('All For Education') movement (2008).

9 Censo Escolar (2018)

10 See Leithwood et al., 2004.

thereof--in the program. Social programs are not often evaluated using randomized experiments in Brazil; however, the approach has long been used by researchers across the world.

Randomized experiments--often referred to as randomized control trials-- are considered the 'gold standard' for understanding the causal impact of social policies and programs. Successful randomization to a given treatment or program solves the problem of selection bias. That is, if participants are randomly assigned two groups--one that receives treatment and one that does not--then researchers can expect that the two groups to be similar, on average, prior to treatment. Without randomization, differences between who participates and does not participate are susceptible to selection bias, or the differential effects of a program based on who decides to participate in it. Randomized control trials reduce selection bias, allowing researchers to measure the true impact of the intervention. As a result, any differences in outcomes between those that receive treatment and those that do not can be interpreted as the average causal effect. In other words, we can be confident that any differences are due to the program itself, not the characteristics of the schools and students who participated in the program.

Randomized control trials are common in medical research. The first medical study to use experimental design was published in 1948 by British researcher Austin Bradford Hill. His work identified the effect of using an antibiotic in the treatment of tuberculosis.[11] Today, new medicine is often evaluated using randomized control trials. Using this methodology, one group receives the treatment and another group receives a placebo, and researchers carefully observe differences between the two groups. This approach has resulted in various recommendations from health organizations that have saved millions of lives. The use of randomized control trials has extended beyond medicine into the social sciences in the few decades. In 2019, economists Esther Duflo, Abhijit Banerjee, and Michael Krem received the Nobel Prize for Economics for their use of the method to research health and education in poor and vulnerable communities.

Instituto Unibanco and the partner states participating in Jovem de Futuro choose to use a randomized experimental approach to measure the causal effect of the program on student achievement. Initially, some stakeholders were hesitant for schools to do this, because they worried that schools that were not selected to participate in the program would be disadvantaged in some ways. Some argued that all schools should participate in the program instead of creating the treatment and control groups necessary to run a randomized control trial. However, without proper randomization of program

11 Bothwell et al., 2016.

participation, researchers cannot fully measure the impact of a program. As discussed above, randomization helps separate the actual impact of the program from other factors that may be correlated (but not caused by) the intervention. For example, participating schools may be more prone to improvement, even without the program. For this reason, stakeholders decided that stronger evidence was necessary to prove the effectiveness of Jovem de Futuro. By choosing an experimental approach, researchers would be able to determine the true effectiveness of the program before expanding it to more schools.

Rigorous impact evaluations require researchers to distinguish between correlational relationships and causal effects. To understand the difference between correlation and causation, it is worth examining a hypothetical example. Suppose that a local authority decides to change the textbooks in its schools. Four years later, the schools have higher average exam results. The introduction of new textbooks and the evidence of improved student achievement are certainly correlated since they occurred within the same time period. However, just because they are correlated does not prove that the introduction of new textbooks *caused* better student learning. There may be something else happening that influenced these exam results. For example, there could have been a reduction in the number of pupils per class, an influx of children from the private system to the public one (or vice-versa), or a change in the school principals' or teachers' pedagogical choices. In other words, we have no way of knowing exactly what caused exam scores to increase without a proper randomized experiment. In fact, it is possible the textbooks chosen by local officials were less effective than the ones that they replaced and that other factors caused students to improve their exam performance.

In this hypothetical example, how can we evaluate the impact of the introduction of the new textbooks on student achievement? The best way to measure the impact of the textbooks is to utilize a randomized experiment. Using this approach, individuals are randomly assigned to two groups--one that will receive new textbooks and one that will not. Successful randomization guarantees that the groups will have similar profiles, and that each group differs only in the sense that one of them will receive the new books and the other will not. If the group of students that received the new learning materials experiences an improvement in its scores whilst the other group experiences either a decrease, a stagnation, or an improvement that is less than the improvement of the group with new textbooks, it could be said that the introduction of the new books had a positive impact on those students that received them.[12] In other words, the average difference

12 The recognition of a positive impact could even occur in a scenario in which a decline in performance is noted. If both the groups compared were to record worse results, but the group benefiting from that specific policy showed a less intense drop in performance, a rigorous evaluation could suggest that such policy was positive since it prevented the decline in performance being even worse.

between the two groups is the actual, causal effect of using the textbooks. This evaluation methodology provides more than a simple correlation; it proves the actual effect of a specific policy or program.

Instituto Unibanco and the partner states choose to use this approach to evaluate Jovem de Futuro because of the power of randomized experimental design to isolate the true impact of a program or policy. A more detailed explanation as to how the experimental evaluation was conducted can be found in later sections. The next section provides a brief explanation of the program and its history.

# WHAT IS THE PROGRAM?

The Jovem de Futuro program began in 2008 and has been continually improved over time. Since its inception, the program has focused on strengthening the leadership and management abilities of school principals and pedagogical coordinators by providing them with data, indicators, targets, processes, training, assistance and various materials to improve student outcomes. The objective of the program has always been to increase the number of years students are enrolled in school, high school graduation rates, and student academic achievement.

The underlying belief motivating Jovem de Futuro is that educators and school leaders have enormous potential. However, many programs and policies that aim to improve education do not yield meaningful results because of implementation challenges within schools. As a result, the identification of best practices and dissemination of knowledge is limited. If policymakers and practitioners are to achieve meaningful results, they must work together more collaboratively across multiple levels within a state--departments of education, regional offices, and school--during implementation.

As part of Jovem de Futuro, Instituto Unibanco offered training, technical assistance, governance, mobilization, and evaluation to participating states--free of charge--to enable them to work together more effectively. Instituto Unibanco also worked with school principals and pedagogical coordinators, supporting them to reflect on their leadership practices. Institute staff created a framework for reflection and action. This method--called the 'Management Circuit'--is inspired by another method known as 'PDCA', which refers to the 'Planning, Doing, Checking and Acting' of the actions.[11]

11 *The PDCA (Plan, Do, Check, and Act) became popular in the post-war years, based upon the work of W. Edwards Deming.*

The Management Circuit lays out a series of cycles made up of six stages: agreement on targets, planning, execution, assessment of results, sharing of practices, and course correction (re-planning). The Circuit promotes program alignment and coherence in schools by integrating the three most important levels of a state system - the state department of education, its regional offices, and the schools. The circuit ensures that the three levels are aligned to the same calendar, data-driven actions, and protocols.

## 1 Alignment of the 3 administrative levels

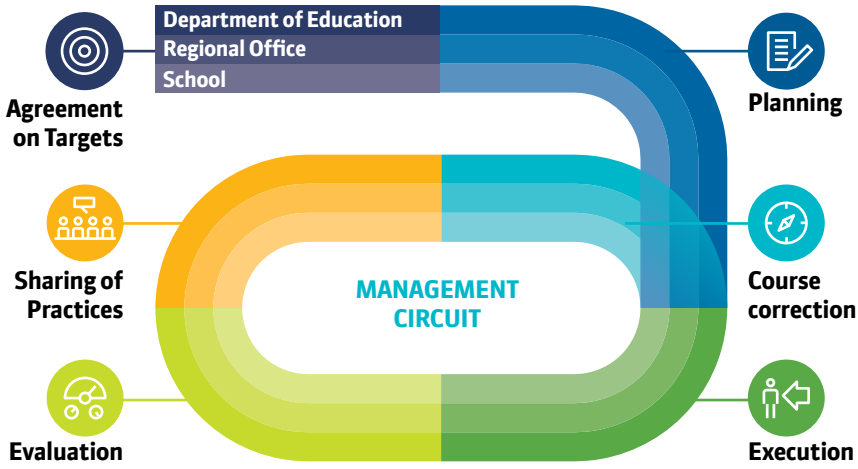



**Diagram 1** – Operation of the Management Circuit in the state systems

The first stage of the Management Circuit involves the agreement on targets. During this stage, stakeholders from the Department of Education and Instituto Unibanco define the educational improvement outcomes that will be achieved over the course of four years. These targets are based on IDEB, the index developed by MEC (Ministry of Education) that measures student promotion rate and achievement in Mathematics and Portuguese Language tests. The four-year program period corresponds to the political cycle of one administration, and the targets for the time period are broken down into four annual goals. This guarantees that the network progresses incrementally towards that overall goals of the program.

This agreement generates a cascade effect: each school assumes a commitment to a specific target. The lowest-performing schools are

considered 'priority' schools They receive greater attention from the regional offices and the department of education to enable them to advance more quickly and reduce the distance that separates them from the rest of the network. Each regional education office also has its own target, calculated using the average of the targets of the schools it is responsible for coordinating.

Once a collective commitment to these objectives has been established, all three levels of educational governance--the department of education, regional offices and schools--move on to the second stage. This stage is focused on planning the actions necessary to meet the targets from the first stage. Stakeholders create action plans that detail the activities to be executed over specific time periods, the results expected, and the people responsible for these activities and results.

Throughout the Circuit, the teams from the regional offices and the departments of education receive technical support and training from the Instituto Unibanco. This support is guided by protocols focused on the activities and indicators that should be monitored. The use of protocols, however, does not mean that the solutions to the problems are imposed on the schools. Quite the contrary. As the aim of the program is to develop the ability to learn through practice, it is the management teams themselves that--at each school and in dialog with the school community--diagnose problems and decide on solutions to solve them.

A management team and school community may, for example, conclude that one of the main challenges to student achievement is the chronic absenteeism. With this in mind, the actions designed to address this problem are considered collectively. Other schools simultaneously perform the same exercise but may arrive at different diagnoses and solutions.

In complex networks with a large number of schools, stakeholders must consider how to effectively coordinate action, monitoring, and support the schools. In the Jovem de Futuro program, support is provided by supervisors who are experts employed by the regional education offices. Supervisors typically visit schools biweekly, but priority schools receive weekly visits.

Jovem de Futuro operates under the assumption that schools need support from all levels of education governance--local, regional, and state. Once schools have identified problems and proposed actions to improve the quality of teaching and learning, regional education offices develop their action plans to support the work in schools. Finally, state departments of education create similarly aligned plans. The Management Circuit encourages the network to work as a coherent whole, with all levels taking responsibility to reach the objectives and plans established in the initial stages.

Once the planning stage is complete, schools, regional offices, and state departments of education must execute on the proposed actions for which they are responsible. Supervisors help schools to monitor their actions and identify any immediate adjustments they can make. This is different from the evaluation stage of the Management Circuit and critically important to successful implementation.

'Formal evaluation' takes place in the next stage of the Circuit. At this point, the entire network pauses their work and collectively reflects on what has been achieved since creating and implementing the plan. This is a bottom-up process; the evaluation is performed primarily in schools, then in the regional offices, and finally in the state departments of education. For example, a school that has identified that the main problem to be addressed is the high rate of student absenteeism will evaluate if absenteeism has decreased and if student achievement has increased as a result. This evaluation is then shared with the regional education office, which, in turn, shares the general evaluation of the actions under its responsibility with the department of education.

The 'sharing practices' stage of the Management Circuit is critical to improving the knowledge of practitioners and policymakers. During this stage, the management teams from all the schools meet at their respective regional offices to share their experiences based upon their mistakes and successes of their work. In the example of the school working to address absenteeism, leaders might share successful practices. Or, if the results have not been satisfactory, this is the opportunity to listen to other schools that have faced similar challenges and determine a new course of action. This process is also undertaken between the regional offices. In both settings, the principal objective of sharing is to improve the policy and practice of stakeholders.

The final stage of the Management Circuit, the correction of routes, encourages stakeholders to revise action plans based on their experiences and the experiences of other schools and offices in their network. Once their plans are revised, the Circuit is restarted and a new cycle of execution, evaluation, and sharing practices begins. Restarting the Circuit ensures that implementation continues even during challenging or unpredictable situations, and that actions are continuously evaluated and shared within the network.

## 1.1 THE THREE GENERATIONS OF THE PROGRAM

Just as the Jovem de Futuro promotes continuous improvement in schools, the program has also improved over time. There have been three significant changes since it began. As a result, we refer to

three 'generations' of the program: the first occurred between 2008 and 2015, the second occured between 2012 and 2015, and the third began in 2015. A transition process was conducted between each generation.

Jovem de Futuro was still in pilot form during the first generation. This initial form of the program involved 197 schools in four states (Minas Gerais, Rio Grande do Sul, Rio de Janeiro and São Paulo), and much of the programming and support was done by employees of the Instituto Unibanco. Although partnerships with state departments of education were established during this time, most of the work was done directly in schools. Instituto Unibanco provided training and technical assistance to schools for the creation, execution and monitoring of a strategic improvement plan. Stakeholders agreed upon a number of ambitious targets that sought to increase learning, reduce the dropout rate, and ensure students were meeting benchmarks for learning before leaving school.[12] These targets were broken down into seven goals,[13] and organized into three central foundations - students, teachers, and school leadership. These goals guided schools' action plans. Progress towards the goals was monitored using external evaluations and data from the schools. The cycle of change during the first generation lasted three years.

To support this improvement process, employees from the Instituto Unibanco visited schools weekly. During these visits, they worked with their assigned mentees. In addition, pedagogical resources and tools designed to mobilize community participation were offered to the teachers. To encourage the schools' adhesion to the new management process, financial resources were transferred directly from the Institute to the schools.[14] These resources could be used to reward teachers and students, improve the infrastructure, better train professionals, or create support funds for pedagogical projects and activities for students.

The success of this pilot created an opportunity to expand the scale and sustainability of the program. Thus began the second generation of the Jovem de Futuro.

12 The cut-off points that define the learning standards or levels were defined by the states themselves, in accordance with their curricula and the local expectations for learning for each year and stage of learning. For the first generation, the Jovem de Futuro database does not contain any information on the local choices and, as such, the cutoff points established for the scale, created in 2006 by the technical commission of the '*Todos pela Educação*' movement and by the researcher José Francisco Soares (UFMG), ex-president of Inep, were adopted.

13 The seven results were: students' learning assessed by Portuguese Language and Mathematics tests; students with a high attendance rate; development of socio-economic-environmental skills; reduction in the absenteeism amongst teachers; new pedagogical practices; improvement of the school infrastructure; and management focused on results.

14 Each participating unit received an annual sum of R$100.00 per student enrolled in regular secondary education.

## 2 The Jovem de Futuro Program and its three generations

### OBJECTIVES

**To guarantee learning, continuity in school and graduation from high school**

### HOW IT WORKS

**Management program**
Cycles of change, execution and monitoring of the plans, with revised planning of the actions

**Focus on pedagogy**
Direction of efforts for the improvement of teaching and learning

**Performance targets**
Improvement of learning and certifications, with a reduction in inequalities

**Management Group**
Collective work performed by the school principal, pedagogical coordinator, teachers and students

**Supervisors**
These individuals provide advice and support for the training of the management group, whilst also monitoring actions

**Actions and Resources**
Involving mobilization, training, technical assistance, governance* and information systems

*since 2nd generation

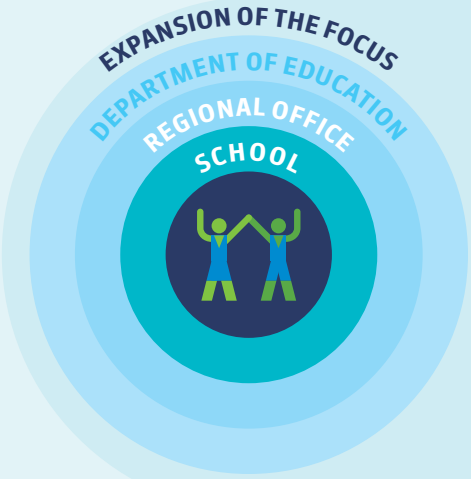

### THE FIRST GENERATION

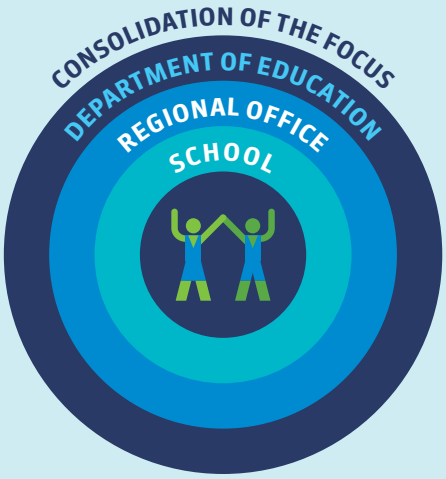
**The Instituto Unibanco works directly in the schools.** The objective, maintained across all generations of the program, is to get management more focused on learning and the students' continuity in high school, whilst also reducing inequalities.

One of the mechanisms proposed to encourage adhesion has been the direct transfer of financial resources from the Instituto Unibanco to each school.

The **cycle of change**, implemented in each school, runs over a course of **three years**.

**197 schools**
in
**4 states**
(MG, RS, RJ and SP)



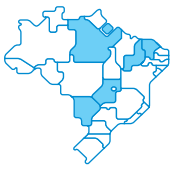

### THE SECOND GENERATION

This marks the shift of a school-based project to a system-wide policy, implemented by the Department of Education, with a view to scale and sustainability.

**The State starts to assume a central role**, making Department supervisors and technicians available. The Instituto Unibanco designs, supports and monitors the implementation. The financial incentive is now guaranteed by the '*Ensino Médio Inovador*' ('ProEMI'/'Pioneering Secondary Education') federal program.

The **cycle of change**, implemented in the school, becomes **annual**.

**2,166 schools**
in
**5 states**
(CE, GO, MS, PA and PI)



### THE THIRD GENERATION

This consolidates the transformation of the Program into an education network policy.

**Management for continuous improvement is introduced**, reinforcing the focus on the student and on pedagogic management, as well as the process of learning through practice. The work becomes systemic and the financial transfers cease.

The **cycle of change** implemented in schools, regional offices and the central body is reinforced and becomes **quarterly**.

**3,549 schools**
in
**7 states**
(ES, PI, GO, PA, CE, RN and MG)

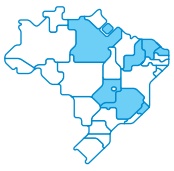


**Diagram 2** – Evolution of the Jovem de Futuro

During its second generation, Jovem de Futuro shifted from a school-based project to a system-wide policy in order to increase the scale and sustainability of the program. Instituto Unibanco designed, supported, and monitored implementation, and leaders in departments of education and regional education offices received training and support as part of the program. School supervisors from regional education offices made monthly visits to an average of 10 schools and supported the cycle of change within schools. These supervisors formed bridges between the schools and the regional education offices, facilitating the flow of information and ensuring schools were part of the network management strategy. This generation included 2,166 schools (87% of all public high school enrollment) in five states: Ceará, Goiás, Mato Grosso do Sul, Pará and Piauí.

The strategic improvement plan implemented by schools underwent two major transformations during the second generation of the program. The first transformation arose from a partnership with the Ministry of Education's '*Programa Ensino Médio Inovador*' ('Pioneering Secondary Education Program' / 'ProEMI'). This program sought to offer evidence-based innovations in Brazilian schools. The vision for this program was similar to the vision of Jovem de Futuro, and, as a result, ProEMI included Jovem de Futuro as part of its federal public policy. ProEMI was implemented throughout the country between 2009 and 2018. It offered schools financial incentives[15] as well as directives for the curriculum improvement of secondary education.[16] Both ProEMI and Jovem de Futuro prioritized "improvement in management" and "curricular reform" in schools. The second major transformation was the introduction of the Management Circuit (discussed in greater detail above), which is based on continuous improvement.

Despite important improvements made during the second generation, a number of challenges remained. Schools needed to be more engaged in order to successfully transform their practices. The bureaucratic processes that were present in the second generation (with the overlapping of the accountability of Jovem de Futuro and ProEMI) also wasted participants' energy and detracted from their focus

15 An average sum of R$70.00 per student enrolled in secondary school was provided, as a means of enabling the implementation of the *projeto de redesenho curricular* (*'curricular redesign project' / 'PRC'*)

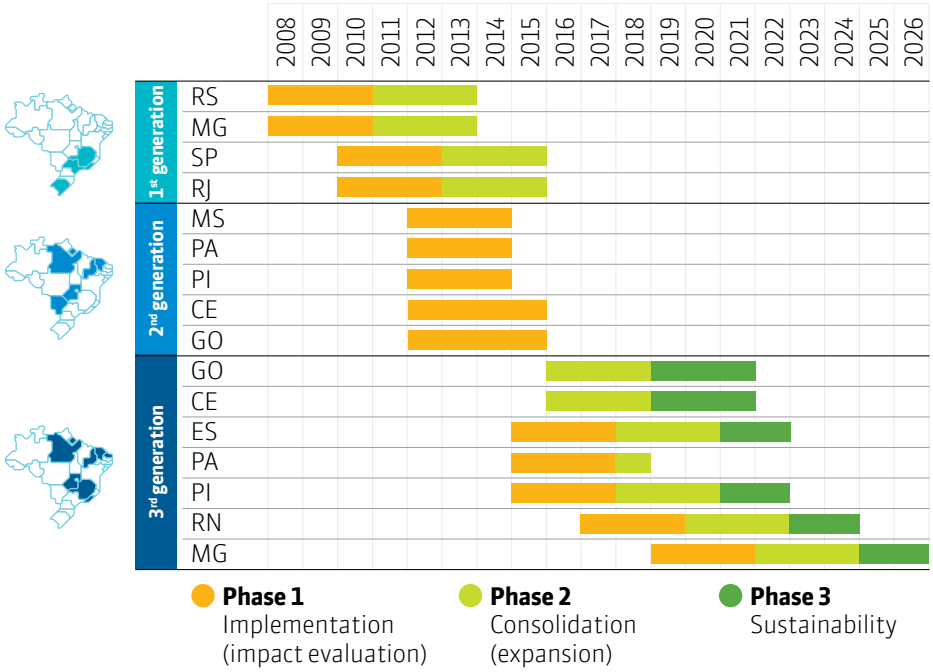
16 The new curriculum was meant to be developed with participation from the entire school community. By doing this, the curriculum would be aligned to the school's political-pedagogic project and respect the following parameters: an increase of the minimum annual workload to 3,000 hours; a guarantee of interdisciplinary actions to be established using four areas of knowledge as a base; and encouragement of the students to sit the '*Exame Nacional do Ensino Médio*' ('National Secondary Education Exam' / 'Enem'). Furthermore, the school needed to respect the pedagogical and curricular fields of actions proposed by the Ministry of Education, referred to as 'micro-fields'.

on students. Additionally, the federal ProEMI program placed large bureaucratic demands on schools with limited funds. As a result, schools were forced to find solutions to these demands without financial support from the federal government.

The third generation of the program aimed to expand the Management Circuit to include regional offices and state departments of education. During this time, the program sought to create coherent plans that aligned all levels of education governance (schools, regional offices, and state departments of education). This was critical because schools have relatively limited autonomy to make hiring, training, and infrastructure improvement decisions. The third generation of the program stopped pursuing curricular reform, and concentrated on school management and instructional leadership instead. The program established educational leadership standards for the first time and expanded to include 3,549 schools in seven states.

Finally, we should consider the fact that, in education, sustainable processes of change take time to establish themselves, since they have to work with deeply entrenched behaviors and cultures. As such, Jovem de Futuro's process of change (in partnership with states and localities) covers an eight-year period (see Diagram 3). During the first three years, the focus is on the dissemination and experimentation of the new management method by all those involved. From the fourth to the sixth year of the partnership, there is an intensification of the transfer of knowledge and technologies, so that, at the end of the period, to ensure sustainable implication. During the final two years of the partnership, the function of the Instituto Unibanco shifts to involve the monitoring of the cultural change and supporting the management innovation processes that the partnership has been aiming for.

## 3 Duration of the 3 generations of the Jovem de Futuro




Source: Hypothetical example. Formulated internally.

**Diagram 3** – Duration of the three generations of the Jovem de Futuro

# PREPARING THE IMPACT ASSESSMENT OF THE JOVEM DE FUTURO

Jovem de Futuro has partnered with 11 Brazilian states over the course of a decade to improve school management. The program aims to ensure better decision making within schools and follows a well-defined continuous improvement cycle. Three distinct generations of the program have emerged over time, each with its unique context, actions, and goals. Thus, we measure the impact of the entire program as well as the impact of the three generations.[17]

This impact evaluation focuses on the program's goals: improving student learning, their continuity at school, and the graduation rate in high schools. We use standardized Portuguese Language and Mathematics tests conducted by each state to measure student learning in the third year of high school. In addition to measuring the average scores, we also measure the number of students who fall within a "critical level of learning" due to extremely low scores. We measure continuity in school and graduation rate by tracking student scores over time. These data are released to the public by INEP, the Ministry of Education's institute of evaluation and statistics.

It is important to note that the unit of analysis for this evaluation is schools and not individual students. Students may transfer to a different school, be held back and repeat a grade, or drop out of

17 The Appendix to this publication provides information on the impacts detected in each partner state and in each generation. Analyses of each specific state, however, will not be developed. The decision has been taken to analyze the generations of the program rather than the states, since larger samples offer more precise estimates. The fluctuations in the impact, when comparing states, are explained above all by statistical variations. They are not, therefore, real variations.

school over the course of the study. Thus, measures reflect the average achievement and enrollment of students at a given point.

It is also important to note that while impact evaluations can measure how much a program impacted specific outcomes, they do not explain how or why this impact occurred. As a result, this impact evaluation does not identify the mechanisms through which positive or negative change occured. While we can say with confidence that Jovem de Futuro had a positive effect on schools, it is not possible to "open the black box" and identify the specific actions that generated the greatest impact.

Measuring the impact of the third generation of the program was especially challenging. During this generation, the program took a systemic approach and integrated the actions of schools, regional offices, and the state department of education. Because the end goal of this work was to improve schools, the evaluation measures the difference between those that did and did not implement the Management Circuit. However, a spillover effect may have occurred. In other words, schools that were not participating in the program may have still benefited from the work being done in their regional offices and state departments of education. The evaluation does not take this into account. Therefore, it is possible that the impacts measured by the evaluation underestimate the true effect of the program.

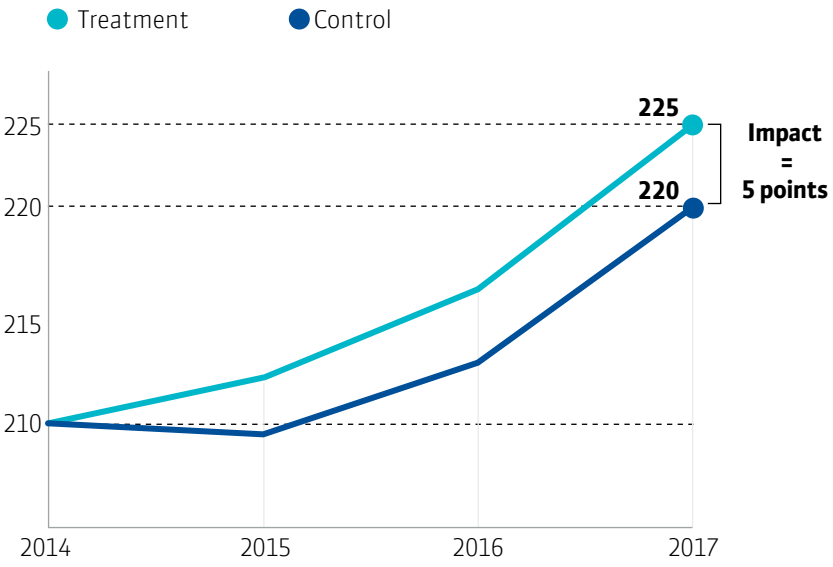
An advantage to using a randomized experimental evaluation is that it is easy to see the results of the program. Let us return to the hypothetical example cited in the introduction to this report. Recall that schools were randomly assigned to use a new textbook. In this example, both groups that receive the new textbook and those that did not are similar in all observable ways *except* that one group is using the new textbook. Let us suppose that the treatment group - that which received the new books -- improved its scores on the mathematics standardized test[18] by 15 points--jumping from 210 points to 225 points over a period of three years. We cannot say that the new textbook improved test scores by 15 points because over time other factors may have caused test scores to increase. Instead, we must compare the treatment group (those randomly selected to use the next textbook) to the control group (those randomly selected to not use the new textbook. As long as randomization has been successful,

18 The standardized state tests commonly use the same scale of the Basic Education Assessment System (Saeb), that brings together a large scale set of external evaluations developed by Inep, intended to offer a diagnosis of Brazil's basic education. For more information, see <http://portal.inep.gov.br/educacao-basica/saeb>. The system was created in order to allow a comparison of the scores achieved by basic education students in the 5th and 9th years of elementary school and the 3rd year of high school, thus demonstrating the evolution of the quality. Specifically for secondary education, the scores for mathematics range from 225 to 475 points. This range, in turn, is divided into nine bands of proficiency, of approximately 25 points each.

the average difference between these two groups is the causal effect of the program. Successful randomization can be checked by checking to see if the groups are similar across multiple observable characteristics. In this hypothetical example where we assume perfect randomization, if the control group improved 10 points, the effect of the textbook is five points.

**1 Conceptual example of the impact measured by experimental evaluation**

Proficiency in mathematics in the covered schools (points on the Saeb scale)




Source: Hypothetical example. Formulated internally.

**Graph 1** – Evolution of the scores in mathematics in a hypothetical case

As noted above, experimental evaluation relies on successful randomization. In the Jovem de Futuro, the randomization process was transparent to participants. All schools received information about the program and the evaluation, and all schools participated in the program. However, schools had staggered implementations. The treatment group schools started in the first year, whilst those in the control group become involved in the beginning of the fourth year.

Before randomizing, the schools were grouped into clusters of two or three schools that had similar observable characteristics.[19] Clustering was important for three reasons. First, clustering ensured that treatment and control schools were balanced across observable characteristics. If the schools had not been grouped, treatment and control groups could have been different, even with perfect randomization. For example, the treated group could have randomly had a higher proportion of large, urban schools, whilst the control group could have randomly had a higher number of small, rural schools. Dividing the schools into clusters ensured more balance between the two groups. Additionally, creating clusters of similar schools meant that at least one school from each profile would be guaranteed to participate in the first year of the program. Finally, this strategy alleviates a common problem in randomized experiments: imbalance due to attrition.

Attrition occurs when participants drop out of the sample over time. For example, schools participating in the program may be closed down or may stop taking part in the state tests that measure improvement over time. When schools drop out of the sample, they can create imbalanced treatment and control groups. If schools leave the sample for reasons that are not connected to the program (or in other words, the exit would have occurred even if the program had not been operating in the state system), it is possible to exclude schools that exit and their corresponding cluster from the analysis. This maintains the integrity of the experiment.[20]

19 Of the 1,160 schools that form part of the impact evaluation of the Jovem de Futuro, around 60% were grouped into pairs and trios. For more information, see Barros and Franco, 2018.

20 Originally, 433 clusters were formed for the purposes of impact evaluation – 84 in the first generation, 101 in the second, and 248 in the third. In the first generation, 31 clusters (37%) were lost, of which 25 were in Rio Grande do Sul state, getting left out of the impact evaluation due to the fact that the state did not have its own Portuguese Language and Mathematics evaluation, meaning exams applied by the Instituto Unibanco were used as a source. The impacts noted proved to be extremely high. Some hypotheses indicated a bias in the treated schools, which had a different motivation for performing the test due to their fear of losing the program, or even the possible dissemination of the items that make up the tests, since the diagnostic tests that formed part of the treatment offered by the program were also used. As such, the evaluation team decided to take a conservative line and exclude this excessive impact from the accumulated estimates. In the first generation, an agreement was made with the schools that any lack of adhesion to the program would lead to the partnership being terminated. Therefore, the other six clusters lost occurred due to the termination of the partnership. One of the cases was due to the state's decision to close a school. In the second generation, fourteen clusters (14%) were lost, and in the third, just eight clusters (3%), due to the closure of schools, or to their non-participation in the state evaluation.

Through until the end of 2019, the impact evaluation of the Jovem de Futuro involved 380 clusters and 1,161 schools in nine states. As a result 380 randomized draws took place to assign schools within each cluster to either the treatment of control groups.

# DOES THE PROGRAM HAVE AN IMPACT?[21]

Evaluators used statistical inference to determine if the Jovem de Futuro program had an impact. They compared schools within each cluster to determine whether, on average, the treated schools (those that had access to the program during the first year of implementation) performed better than the control schools (those that were included in the program during or after the fourth year of implementation). It is important to note that not all treated schools would have shown improvement, even if the program had an impact on average. The question is: in at least how many groups is it necessary to observe that the treatment group performed better before it can be safely confirmed that there has been a positive impact?[22]

In the evaluation of the Jovem de Futuro, there are 380 clusters.[23] If the program had no impact, then differences between the treatment

21 This chapter was developed using as a basis the results and documents that have been produced by Ricardo Paes de Barros and his team since 2007. An academic summary of the work was published in the article repository of the World Bank in 2018, and will be published in an article that is currently being finalized. For more information, see Barros et al. (2018) and Barros et al. (in print)

22 In reality, this rationale, typical of statistics, establishes a test that evaluates the null hypothesis, that is, that we cannot conclude that the program has an impact. The ideal would be that the evidence available allows for a rejection of the null hypothesis. In logical terms, rejecting the hypothesis that the program has no impact is the same as confirming that it does have an impact.

23 Despite Pará, Piauí, Goiás and Ceará participating in the second and third generations, the data drawn from the schools in the last two states was used exclusively to estimate the impact of the second generation. In the third generation, the Program started to be fully implemented into all the schools in these states. In Piauí, it was decided that the second generation should be interrupted before the impact evaluation had been concluded. Because of this, a new draw was performed in the third generation and the evaluation was restarted. The data from the schools in Piaui is therefore being used only to estimate the impact of the third generation. In Pará, the tiers in the second generation were followed through fully in the implementation stage, and as such it was possible to use them in the corresponding evaluation. However, a new draw was performed in this state in the third generation. As a result, Pará appears in the second and third generation estimates with different tiers.

and control schools in each group would be random. That is, in half of the clusters (190), treatment schools would improve more than the control schools, whilst in the other half, the inverse would occur. Significant imbalances in this "50/50" division could suggest a positive or a negative impact.

How many cases above or below 190 clusters will allow us to safely confirm that the impact is positive or negative? The response will depend upon the margin of error around an estimate. In the evaluation of the Jovem de Futuro, we performed a hypothesis test in which we check to see if the null hypothesis (that the program had no impact is true). The test indicates whether there is sufficient evidence to accept this null hypothesis or reject it. The significance level defined expresses how confident the evaluator is that they are not rejecting the null hypothesis when the null hypothesis is indeed true. For example, with a significance level of 10%, the result of the test is reliable in 90% of the cases. In other words, in only 10 percent of the cases would researchers think that the program has had an impact, whilst in reality it has not. With this test design, for a sample of 380 clusters, it is necessary that in 207 of them the treated schools have performed better in order to confirm that there has been a positive impact. Likewise, if the treated schools have done better in less than 174 clusters, then the program has had a negative impact.[24]

Graph 2 summarizes the number of clusters in which the treatment group performed better than the control group on the previously determined impact indicators: average scores in Portuguese Language and mathematics at the end of high school; percentage of students at critical level in relation to their scores in Portuguese Language and Mathematics at the end of high school; and the pass rate in all years of high school.

24 These limits were defined using the hypothesis that, when the program has had no impact, each one of the 380 tiers is subject to the outcome of a coin toss, in which there is a 50% chance of the treated school performing better than the control school and vice-versa, establishing a binomial format for the distribution.

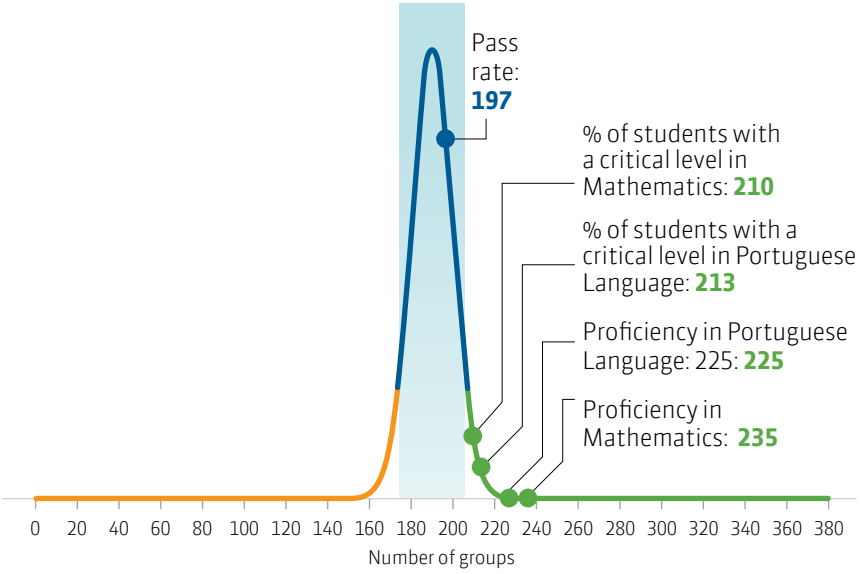
## Where the Jovem de Futuro has had the most impact

Distribution of the number of groups in which the treatment has evolved more than the control of the schools in the Jovem de Futuro Program – 1st, 2nd, 3rd Generation

**Negative consequences** (groups 0 to 174) The majority of groups of schools in these areas point to the program having had a negative impact

**Area of uncertainty in relation to results** (groups 175 to 206)

**Positive consequences** (groups 207 to 381) The fact that there are more groups of schools in this area suggests that the program has had an impact




In relation to the **Pass Rate**, for example, the result falls within the area of statistical **uncertainty**

In relation to the learning in **Portuguese Language and Mathematics** and in the percentage of students at a critical level, we can be **certain that there has been a positive** impact

Source: created using the article published by Barros et al. (2018)

**Graph 2** – Distribution of the number of clusters in the three generations of the Program

Had the program yielded no result, we would expect that half the clusters would yield higher performing treatment schools and half would yield lower performing treatment schools. However, we see that treatment schools have improved Portuguese Language and Mathematics scores (average proficiencies) in more than half of the clusters. We also see a reduction in the percentage of students at the critical learning levels (in both subjects). However, less than half of treatment schools had improved pass rates.

Looking at the green area of Graph 2, the chances of observing so many clusters with a positive impact not arising from the program are low. For example, for the percentage of students at the critical learning level in mathematics, this chance is approximately one in every 2,000 cases.[25] Confirming that the program has an impact, however, is only part of the answer. The extent of this contribution is still unknown.

25 P-values are statistical parameters used to certify that the results found in the theories are not random. They demonstrate the probability of observing a situation that appears to have a positive impact without the measured effect having been generated by the program. The lower the p-value, the lower the probability of this "mistake" occurring. In the evaluation of the Jovem de Futuro, a cut-off point of 10% was chosen for the p-value. This means that estimates with a p-value higher than 10% will not be considered significant. For those tests that seek to prove the existence of the impact, the p-values referring to the estimates for the Portuguese Language score, the mathematics score, and the percentage of students at the critical learning level in mathematics (estimates in the green zone of Graph 2) are, respectively, 0.8%, 0.4% and 0.0%. The probability of observing better performance levels amongst the treated schools and this not having been impacted by the program is, respectively, 1 in 128 cases, 1 in 282 cases, and 1 in 2,141 cases.

# MAGNITUDE OF THE IMPACT

The experimental evaluation considerably simplifies the calculation of the magnitude of the impact. To calculate the magnitude of a program's impact, researchers subtract the average performance of the treatment group from the average performance of the control group. Researchers also calculate a margin of error around these estimates. This margin of error reflects the lower and upper bounds of the estimate. The estimate together with the margin of error summarizes the difference between treatment and control in the population at a given significance level. In other words, the estimate and margin of error help us understand the actual difference between participating in the program (or not) if we were to repeat this experiment. Thus, we can infer the causal impact of the program from our sample.

Table 1 shows the estimates of the impact of the program and the margin of error as well as the p-value, the statistical measurement used to understand whether the results obtained are the result of random fluctuations. The closer the p-value comes to zero, the more confident we are that we have not attributed an impact of the program where there actually is no effect

Nearly all the Jovem de Futuro impact estimates have very low p-values. The majority of these p-values are close to zero, repeating what we have already seen in the previous section: that the chances that we have observed an impact where there actually is no program effect is very low. The chance of seeing effects of this magnitude, or even greater, without the advantage coming from the Jovem de Futuro, is 1 in 1,000 or 2,000 for the scores in Portuguese Language and mathematics. The chance that we have found a false positive effect on pass rate is slightly higher. Given the p-value is 4%, we would expect that if we were to repeat this experiment there is a 1 in 25 chance that there is no effect.

## Jovem de Futuro impact estimates – 1st, 2nd and 3rd generations

| Variables | | Impact after 3 years of JF | | |
|---|---|---|---|---|
| | | Impact (points on the scale) | Margin of error | P-value (%) |
| **Proficiency** | Portuguese Language | **4.4** | 1.6 | 0.0 |
| | Mathematics | **4.8** | 1.3 | 0.0 |
| **Pass Rate** | | **1.4** | 1.1 | 4.1 |
| **% of students at a critical level of proficiency** | Portuguese Language | **-1.0** | 0.3 | 0.0 |
| | Mathematics | **-4.2** | 1.1 | 0.0 |
| **% of students at sufficient or advanced levels of proficiency** | Portuguese Language | **2.9** | 0.8 | 0.0 |
| | Mathematics | **0.9** | 0.5 | 1.0 |

Source: created using the article published by Barros et al. (2018)

**Table 1**

Considering the impact evaluations in all the three generations of the program, in the schools served by the Jovem de Futuro, the impact on the average score in Portuguese Language was 4.4 points. In mathematics, the impact was of 4.8 points. The margin of error is of up to 1.6 points above or below. In relation to school flow, the treated schools saw an average increase in the high school pass rate of 1.4 percent in relation to control schools, which is very close to the margin of error (1.1 point above or below). As we will see in details in later sections, there are significant differences on this number when each of the three generations of the program is considered separately.

As has been mentioned, the Jovem de Futuro also aims to play a part in reducing the percentage of students with critically low levels of learning. The estimated effects in Table 1 suggest that the program reduces the number of students who are performing at these low levels. In mathematics, the treated schools have reduced the proportion of students performing at low levels by 4.2 points in relation to the control schools. The margin of error is 1.1 points above or below. Furthermore, there has been a positive impact amongst the schools' best students (those with learning levels considered to be sufficient or advanced). In

Portuguese Language, once again, the effects were positive, although admittedly more modest.[26]

## 4.1 IMPORTANCE OF THE IMPACT WITHIN THE CONTEXT OF THE BRAZILIAN EXPERIENCE

The quality of Brazilian public secondary education is in a state of stagnation. Compared to the amount of learning that takes place in elementary school, high school students make fewer academic gains. Typically, a student who concluded state secondary school in 2017 increased their Portuguese Language knowledge by 17 points and by 13 points in mathematics, considering the SAEB scale.[27] The low performance rate in Brazilian secondary education has persisted over many years. According to Graph 3, the average score in mathematics in 2013 was approximately 260 points--the same as it was in 2005. In Portuguese Language, exam scores improved by less than 10 points, or less than half the amount of improvement in the primary years, from 2005 to 2013. Year after year, primary education has improved, but secondary education has not. This makes the impact of the Jovem de Futuro even more important.

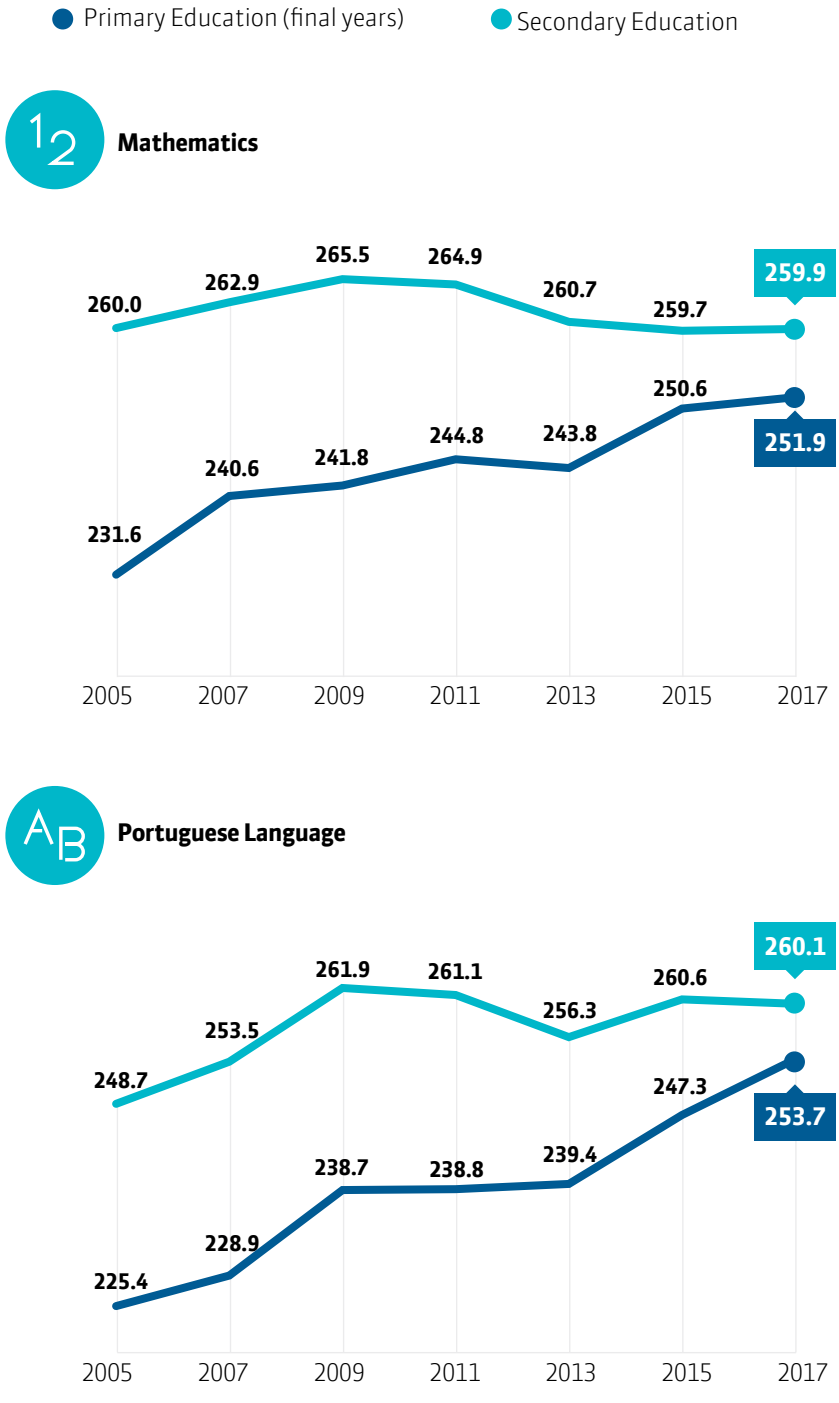
Given these small gains, the five point effect from the Jovem de Futuro corresponds to 26% of what a student typically learns in Portuguese Language during the whole time spent in high school, and 37% of what is learned in mathematics. The effect of the program is equivalent to an additional year of high school mathematics instruction. This is a substantial result.

26 In an article that analyzes the impact of the adoption of good management practices in U.S. public schools, the Harvard-based economist Roland Fryer, highlights that other studies that gage the impact of different interventions designed to improve learning also note better results in mathematics. The author suggests two hypotheses for this phenomenon.First, according to research performed in development psychology, the critical period for the development of language is during childhood. For more complex cognitive functions, the period extends through to adolescence. In the other, learning related to reading is influenced by the context outside the classroom - especially that of the family and community - which could explain why students who speak 'non-standard English' at home or in their community do not perform so strongly. Both explain the estimate of a greater impact for mathematics than for languages. See Fryer, 2014.

27 The minimum score on this exam, that covers from primary untill high school, is 0 points and the maximum score is 425 points in Portuguese and 475 points in math Considering that those who concluded the 3rd year of public high school in 2017 were studying in the 9th year of elementary school in 2014. Since the Saeb exams are given in odd-numbered years, we drew the average of the results from the 9th year students in 2013 and 2015.

## Evolution of the proficiency of the state public Secondary Education and Primary Education system (final years)

Brazil – Saeb scale



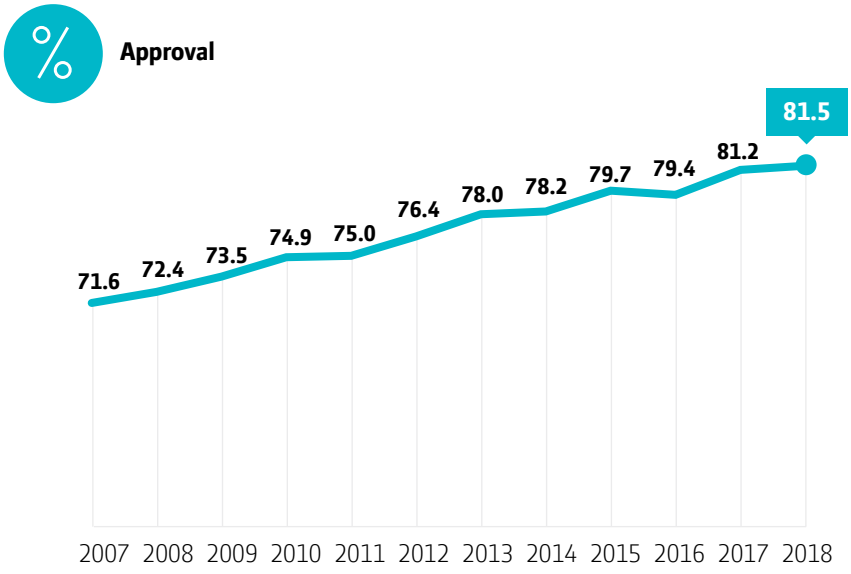

Source: Saeb

**Graph 3** – SAEB scores (2005 - 2017)

In contrast to test scores, the secondary school pass rate has substantially improved since 2007. The average pass rate improved by nearly 10 percentage points over the period, or in other words, by 0.9 percentage points per year. As the impact of the Jovem de Futuro involves three years of intervention, three years of "natural" improvement over this decade would mean an increase in the pass rate of 2.7 percentage points. Jovem de Futuro increased the school's pass rate by an average of 1.4 percentage points. This rate of improvement represents a 50% average increase for participating schools.

**Evolution of the pass rate in the state Secondary Education system**

In %




Source: INEP

**Graph 4** – Secondary school pass rates (2007-2018)

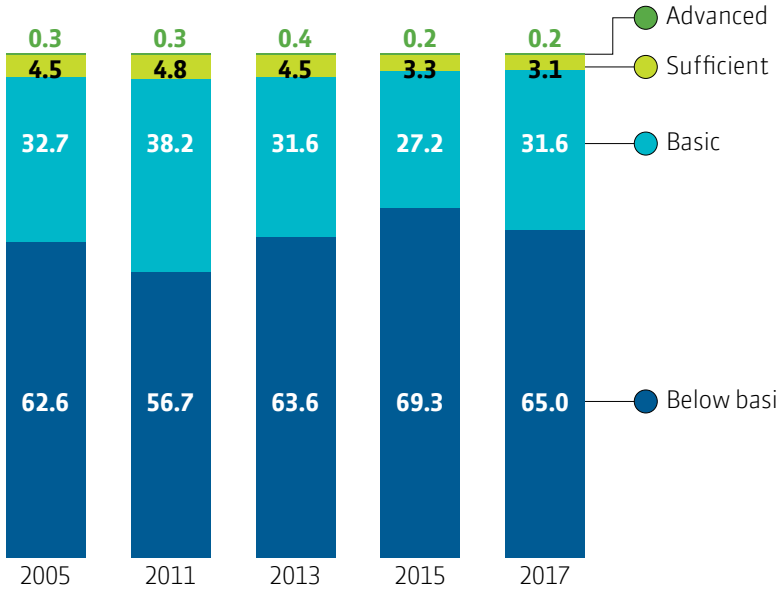
Finally, the historical trends of Brazilian students at the most critical[28] level of learning is an extremely serious matter. Between 2005 and 2017, Brazil went from bad to worse in this area. The percentage of students with mathematics scores at this critical

28 The more critical level should be understood as being the percentage of students who have a learning level below 275 SAEB scale points in mathematics, and 250 points in Portuguese Language. This cut-off point is the same used by the researcher Francisco Soares, ex-president of INEP, the one responsible for designing the 'SARESP' (*Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo* / 'São Paulo State School Performance Evaluation System'), and one of those who developed the expert report that established the levels of learning considered to be sufficient or insufficient.

level rose from 62.6% to 65%. In contrast, participation in Jovem de Futuro decreased the number of students performing below these critical levels (see Figure 5).

**Evolution of the percentage of students by level of performance in Mathematics**

In %

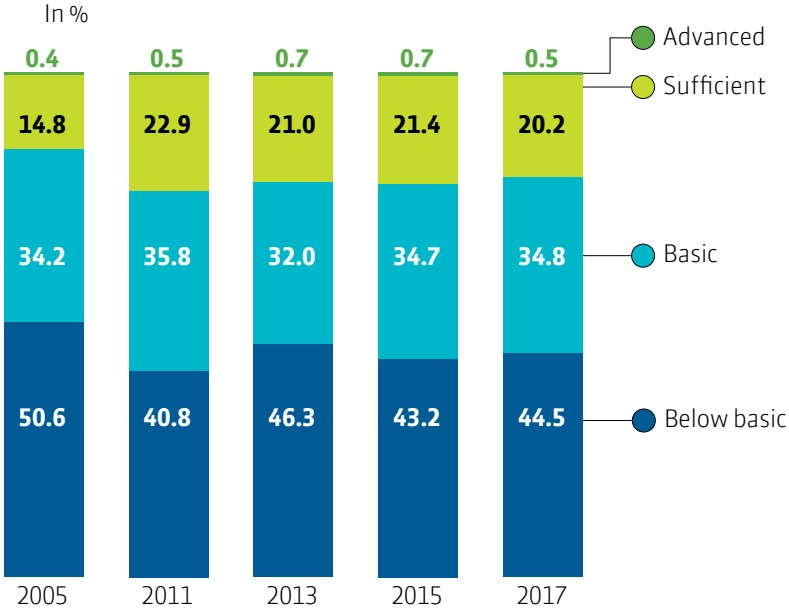



Source: developed internally based upon SAEB results. The four levels were established using an interpretation of the scale performed by Franciso Soares as a base.

**Graph 5** – Distribution of learning standards in mathematics (2005 - 2017)

## Evolution of the percentage of students by level of performance in Portuguese Language

In %




Source: developed internally based upon SAEB results. The four levels were established using an interpretation of the scale performed by Franciso Soares as a base.

**Graph 6** – Distribution of learning standards in Portuguese Language (2005 - 2017)

## 4.2 INTERNATIONAL COMPARISONS

Next, we consider whether the impact of Jovem de Futuro is comparable to the impact of school management programs in other countries. An increase of five points on the SAEB scale in mathematics learning does not necessarily mean the same thing on an international scale. The key is to convert the values into standard deviations, a gage used by statisticians to measure the spread of a distribution in relation to its average.

The scores on Portuguese Language and mathematics exams reported in previous sections are in reference to the SAEB scale. The minimum score on this exam is 0 points and the maximum score is 425 points in Portuguese and 475 points in math.[29] The five points gain in mathematics proficiency that resulted from participation in Jovem de Futuro represents a 0.10 standard deviation increase in performance.

To understand the magnitude of this effect, we examine other international experimental evaluations concerned with school management and other variables that might impact students' learning. A more in-depth study of other rigorous studies using a similar approach is under way. One limitation to this exercise is that the scientific articles provide little information on the design of the interventions, making it difficult to compare the effects.

An important and widely respected study about school leadership was conducted by Roland Fryer, of the Department of Economics at Harvard University. In 2017, the economist demonstrated that the students in schools where the principals participated in a high-quality leadership training program demonstrated improved learning in a range of subjects, including English, arts, mathematics, social sciences and natural sciences. Results suggest that these effects were as large as 0.10 standard deviations.[30] The two-year program being evaluated offered school principals 300 hours of training in school planning, data use, class observation and coaching. In an earlier experimental study, from 2014, Fryer compared the performance of a group of public schools randomly chosen to receive training to practices from high-performing 'charter schools'[31] such as tutoring, a greater amount

29 The distribution of scores for Portuguese Language and mathematics for secondary education, in the public system, was considered.

30 The experimental evaluation was conducted in public schools in Houston, in the United States. The effects reported using high-stakes test scores were of 0.10 standard deviations, and the effects reported using low-stakes test scores were of almost 0.20 standard deviations more in relation to performance. For more details, see Fryer, 2017.

31 'Charter schools' are privately-run, public schools that receive public resources and provide free education to the students. The study compared traditional public schools in Houston that adopted practices that were common in charter schools with public schools that did not adopt these practices. None of the schools in the experiment became a charter school during the period of study.

of time dedicated to classroom teaching, and a proliferation of high expectations. There was a statistically significant 0.15 standard deviation impact of learning in mathematics. In reading, however, the effect was not statistically significant.[32]

A meta-analysis published in 2007 by Carolyn Hill, of the Georgetown Public Policy Institute, and co-authors in partnership with the MDRC (Manpower Demonstration Research Corporation) provides greater international context. This study brought examined evaluations of educational interventions focused on improving teaching and learning. Taking into consideration only those experiments that involve secondary education, the average impact found in 43 evaluations was of 0.27 standard deviations. Although the publication does not go into detail about the type of intervention that has the greatest impact on the students' learning, there is evidence that actions focused on the teachers are the most effective. This confirms work by Eric Hanushek, of the University of Stanford. According to him, teacher quality has the largest impact on student achievement with highly effective teachers improving student learning by 0.10 to 0.20 standard deviations. In comparison, the impact of the Jovem de Futuro over a period of three years 0.10 standard deviations--about one third that of the effect of teachers according to Hanushek.[33]

In general, more than one third of the experimental evaluations included in the meta-analysis organized by Hill et al. (2007) showed impacts that were less than the effect of Jovem de Futuro. According to the meta-analysis conducted by the Brazilian economists Paes de Barros and Portela,[34] interventions that reduced the number of students in the classroom or replaced novice teachers with those who affected student learning to a similar extent as Jovem de Futuro. Based on existing evidence, we conclude that the impact of Jovem de Futuro is similar to other programs that seek to improve school management. That is, the program has a substantial impact, but not quite as large as the impact of teachers on students.

While non-experimental evaluations are not directly comparable to this impact evaluation, they nonetheless describe associations and relationships between interventions and student achievement. In 2015, Professor Nicholas Bloom of Stanford University, together with other co-authors, investigated the relationship between school principals' leadership practices and student mathematics achievement. The

32 This study also included public schools in Houston, in the US, with a public different to the previous. For more details, see Fryer, 2014.

33 See, for example, Hanushek and Rivkin (2006, 2010), that show the effects of a number of different studies on the subject.

34 *Caminhos para melhorar o aprendizado*. Available at: <www.paramelhoraroaprendizado.org.br> Accessed on: May 8, 2020.

authors' study included more than 1,800 high schools in eight countries, including Brazil[35] and found that effective leadership is associated with a 0.24 standard deviation increase in student learning. However in Brazil student learning increased by only 0.10 standard deviations.

In short, effective school leadership and management is important for the transformation of education. The Jovem de Futuro has provided an important contribution to the Brazilian scenario, not only in terms of results, but also for the generation of scientific knowledge on education. The greater impact of one specific initiative does not mean, however, that only one type of intervention should be considered. In general, broader reforms that result in sustainable improvements seek to improve the teaching and learning conditions on various fronts. Each policy carries its own importance.

35 The countries included in the study were: the United Kingdom, Sweden, Canada, the United States, Germany, Italy, Brazil and India. See Bloom et al., 2015.

# WHAT CHANGED IN RELATION TO THE IMPACT OVER THE GENERATIONS?

The first generation of the Jovem de Futuro was a pilot program. The results from this pilot provided evidence that the program would positively impact student achievement and outcomes. In the second generation, the Jovem de Futuro was expended and evolved into a comprehensive strategy for improving entire education systems. During this time, state departments of education led a comprehensive implementation of the program. During the first generation, less than 5% of schools--making up less than 10% of students--in each state participated in the program. In the next generations, these numbers increased substantially. See Table 2 for more details.

## Maximum coverage achieved in the Jovem de Futuro partner states that had concluded the impact evaluation by 2018

| | Partner States | Number of Schools | Percentage of Schools | Percentage of Enrolments |
|---|---|---|---|---|
| First Generation | Minas Gerais | 44 | 2 | 6 |
| | Rio Grande do Sul | 46 | 4 | 8 |
| | Rio de Janeiro | 30 | 3 | 8 |
| | São Paulo | 77 | 2 | 3 |
| | **Total *** | **197** | **2** | **5** |
| Second Generation | Ceará | 439 | 69 | 76 |
| | Goiás | 580 | 93 | 93 |
| | Mato G. do Sul | 271 | 88 | 88 |
| | Pará | 458 | 86 | 93 |
| | Piauí | 418 | 88 | 90 |
| | **Total*** | **2.166** | **84** | **87** |
| Third Generation | Espírito Santo | 237 | 91 | 95 |
| | Piauí | 451 | 91 | 90 |
| | Goiás | 590 | 100 | 96 |
| | Pará | 203 | 33 | 33 |
| | Ceará | 640 | 99 | 99 |
| | **Total*** | **2.121** | **68** | **70** |

* The percentage of schools and enrolments served relates, respectively, to the total number of schools and enrolments achieved by the JF in relation to the total number of Secondary Education schools and enrolments in the participating states.
N.B.: in the third generation, the states of Minas Gerais and Rio Grande do Norte were not included since the prior only started participating in the program in 2019 and the latter was in the 'evaluation window' period.
Source: developed internally

**Table 2**

Despite the expansion between the first and second generation, and the transfer of a significant part of the implementation to the departments of education, the impacts continued with more or less the same pattern, as can be seen in Table 3. In the third generation, the effect on the average scores in Portuguese Language and mathematics dropped slightly, while at the same time-- and for the first time--the program started having an impact on the secondary education pass rates.

**X** **Impact on proficiencies and approval rate per generation**

Impact after 3 years of JF on:

| JF Generations | Portuguese Language | | | Mathematics | | | Pass Rate | | |
|---|---|---|---|---|---|---|---|---|---|
| | Impact | P-value (%) | Margin of error between generations | Impact | P-value (%) | Margin of error between generations | Impact | P-value (%) | Margin of error between generations |
| Average | 4.4 | 0.0 | – | 4.8 | 0.0 | – | 1.4 | 4.1 | – |
| 1st Generation | 5.5 | 3.8 | | 5.3 | 2.7 | | 1.1 | 33.5 | |
| 2nd Generation | 5.6 | 1.3 | 4.4 | 5.8 | 0.7 | 3.6 | -0.8 | 25.9 | 4.1 |
| 3rd Generation | 3.1 | 5.5 | 3.5 | 3.7 | 2.9 | 3.0 | 2.8 | 2.9 | 2.5 |

Fonte: elaborado a partir do artigo de Barros et al. (2018)

**Table 3**

There is evidence that the fluctuations in the impact on the scores between the second and third generations are not important, since they fall within the estimated margins of error.[36] However, the differences of impact on the pass rate led us to believe that the pattern of the program shifted during the third generation. The impact on the pass rate increased 3.6 points, whilst the margin of error of the estimate is of approximately 2.5 points above or below.

The fact that the reduction of the impact on the scores is within the margin of error leads us to doubt whether an actual reduction took place. It is possible that the increase of the impact on the secondary education pass rate was accompanied by a drop in the scores of those students who reached the 3rd year. Ultimately, if the pass rate in each one of the years of high school increased, this means that those students with a more vulnerable profile, who had previously not been graduating the year, were now doing so. Even taking into consideration that their learning was evaluated by the schools as being

36 The standard error of the impact difference between the second and third generations is given by the square root of the sum of the squares of the errors of the impact estimate of each generation. For a statistical significance of 0.10, this standard error of the difference was multiplied by 1.65, in accordance with the normal standard table.

sufficient to pass them, the 3rd year of high school starts to become more heterogeneous, with students from different backgrounds, and this can reduce the scores of the treated schools in comparison to the previous scores.[37] It is interesting to note that the greatest effect observed on the pass rate takes place in the 1st year, which is precisely the year in which there is the highest rate of dropping out of school and truancy.

Table 4 shows the average impact of the program on pass dates in the first three years of high school. In the first year, schools that participated in the program had pass rates that were 3.6 percentage points higher than the control schools. In the second year, the impact was 2.1 percentage points. However, the effect of the program was no longer significant in the second year of the program at the 10% significance level. Therefore, the new pattern of impact as of the third generation seems to move in a direction that raises social inclusion, with a reduction in educational inequalities.

**X Impact of the 3rd generation on the pass rate by year**

In percentage points

| Grade | Average impact | Margin of error (in points) | P-value (%) |
|---|---|---|---|
| 1st year in high school | **3.6** | **2.1** | **4** |
| 2nd year in high school | **2.1** | **1.9** | **8** |
| 3rd year in high school | **1.0** | **1.5** | **18** |

Source: created using the article published by Barros et al. (2018)

**Table 4**

Table 5 provides evidence that the number of students score at the critical level or below decreased as a result of the program. The results suggest that this trend continues across all three generations of the program.

At the other end of the achievement distribution, the percentage of students with sufficient or advanced scores was impacted by Jovem

37 The impact of the socioeconomic level of the families on the learning, verified in tests, is one of the most solid pieces of evidence in educational research. The first study to identify this association was published in the United States in the 1960s, by the sociologist James Coleman (Coleman et al.,1966).

## Impact on the distribution of grades

Impact after 3 years of JF on:

**% of students at a critical level (level 1)**

| JF Generations | Portuguese Language | | | Mathematics | | |
|---|---|---|---|---|---|---|
| | Impact | P-value (%) | Margin of error between generations | Impact | P-value (%) | Margin of error between generations |
| Average | -1.0 | 0.0 | – | -4.2 | 0.0 | – |
| 1st Generation | -5.2 | 2.6 | 2.8 | -4.5 | 3.1 | 2.8 |
| 2nd Generation | -1.0 | 0.3 | 2.1 | -4.3 | 0.2 | 2.4 |
| 3rd Generation | -2.2 | 9.2 | | -4.1 | 2.2 | |

**% of students at the desired level (levels 3 and 4)**

| JF Generations | Portuguese Language | | | Mathematics | | |
|---|---|---|---|---|---|---|
| | Impact | P-value (%) | Margin of error between generations | Impact | P-value (%) | Margin of error between generations |
| Average | 2.9 | 0.0 | – | 0.9 | 1.0 | – |
| 1st Generation | 3.4 | 3.8 | 2.3 | 0.6 | 11.5 | 1.0 |
| 2nd Generation | 3.3 | 0.2 | 1.8 | 1.9 | 0.9 | 1.0 |
| 3rd Generation | 1.4 | 11.0 | | 0.3 | 27.9 | |

**Table 5**

de Futuro (see Table 6). Schools that participated in the program had a higher percentage of students performing at the desired level on exams. In the third generation, this effect was more modest. It is possible that the most significant effects on the group with lower scores are indicative that the third generation was able to reduce the inequality between the students, without flattening the high-end of the distribution of scores, but rather having a greater effect on the less favored.

# NEW QUESTIONS AND THE SCIENTIFIC DEBATE IN EDUCATION

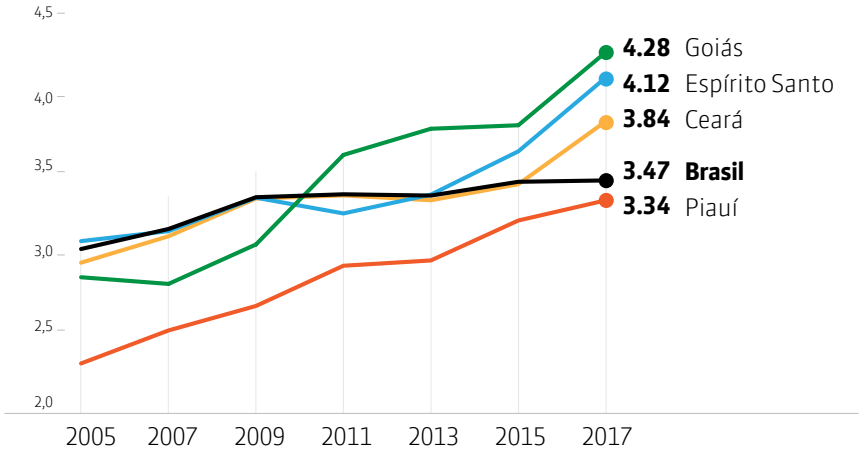
The results from Instituto Unibanco's Jovem de Futuro program confirm that improving school leadership and management is critical to student achievement and completion of secondary education. The success of the program in a variety of contexts--from the northern to southern extremes of the country--highlights the power of the program and adds to the scientific research on educational interventions.

For more than a decade the program has influenced the management strategies of education practitioners and government agencies. As best-practices have spread across schools, regional offices, and state departments of education, it has become more difficult to estimate the true effect of the program. This is because the benefits of the program have improved beyond the original treatment schools. To be clear, this is a good thing. Education systems are improving by adopting leadership and management best-practices, ultimately benefiting students. As schools adopt more effective planning and pedagogical approaches, we have seen a transformation in the education policies and practices across the country, even in schools in the treatment group. This spillover suggests that we may be underestimating the impact of the program.

One way to understand how the program indirectly improved education systems is to look at IDEB scores of states that participated in Jovem de Futuro. It is important to note that we cannot draw causal conclusions about the effect of the program by looking at this data. As a reminder, IDEB stands for the Basic Education Development Index and is the Ministry of Education's most important index of student achievement. It is used throughout the country to monitor advances in the quality of teaching and made up of a combination of the scores in Portuguese Language and mathematics, as well as the pass rates.

It is interesting to note that, in all the states where the program advanced in scale (Ceará, Goiás, Espírito Santo and Piauí), there has been a systematic improvement in the respective scores on the IDEB since the beginning of the partnership. The managers in these states feel that the partnership with the Instituto Unibanco is one of the elements that have helped them achieve this result.[38] [39] However, further evidence from an experimental evaluation would be necessary to support the claim that the program improved educational outcomes at the state level.

### Evolution of the IDEB of the States




| Year of entry into the Program | State | 2nd generation | 3rd generation |
|---|---|---|---|
| | Ceará | 2012 | 2016 |
| | Espírito Santo | - | 2015 |
| | Goiás | 2012 | 2016 |
| | Piauí | 2012 | 2015 |

Source: developed internally using INEP data as a base

**Graph 7** – Evolution of the Ideb in those states where the JF was applied on a large scale

In some states, the Instituto Unibanco performed an exercise in an attempt to estimate the contribution of the program to their improvement on the Ideb. As explained in previous sections, the unit of analysis used at impact evaluations from Jovem de Futuro were

38 Instituto Unibanco, *Aprendizagem em Foco* bulletin, n. 49, Mar. 2019. Available at: <https://www.institutounibanco.org.br/aprendizagem-em-foco/49/>. Accessed on: May 8, 2020.

39 See Henriques and Rocha, 2018.

schools. However, to estimate the contribution of the program at the average IDEB from an entire state, it was necessary to use the IDEB (a federal assessment conducted every two years) and an analogous indicator called IDEB* or "IDEB Star" (administred annually at state level).[40] Assuming that IDEB and IDEB* are entirely comparable, the impacts measured by the experimental evaluation of the third generation indicate that around half the total improvement on the Ideb of Espírito Santo and Piaui[41] would not have occurred without the program. The fact is that the real contribution of the Jovem de Futuro may be even greater, given the improvement in management that is occurring throughout the education system.[42]

The 2019 expansion of Jovem de Futuro in Minas Gerais will allow researchers to better understand the relationship between the program and system-level improvement at scale. Only 44 schools in Minas Gerais were involved in the program during the pilot stage. A decade later, in 2019, more than 2,300 high schools and 47 regional offices participated. This will allow Instituto Unibanco to maintain the historic series of evaluation and, also, investigate two new hypotheses.

First, it will be possible to investigate the impact of regional offices. Until 2018, the number of regional offices in each state was too small to allow statistical comparisons between treatment and control groups. The enormous scale of the Minas Gerais network allowed this type of innovation to be incorporated. It would therefore be possible to understand the true relevance of the management in the regional offices for the improvement of the schools' results, and one more step forward would be taken towards understanding the full effect of Jovem de Futuro.

Second, the expansion of Jovem de Futuro in Minas Gerais leads to important questions about the impact of educational leadership and student outcomes. In Brazilian and international academic writings, a great deal of value is placed on "instructional leadership." This refers to the monitoring of the pedagogical process by the school principals, their support for teachers, and an investment in professional development

40 Ideb* is the generic name given to the annual indicator, analogous to the Ideb, put together for each state using information drawn from the state evaluations and the pass rates, according to the same calculation formula as the national index.

41 Despite Pará having participated in the impact evaluation of the third generation, the coverage of schools remained low. It is therefore not right to include it in this exercise that considers the generalization of the impact. For more information on the level of impact in each state, see the Annex.

42 It is worth noting that three of the Jovem de Futuro partner states (Goiás, Espírito Santo and Ceará) were amongst the four highest positions in the IDEB state secondary education ranking for 2017. Pernambuco, despite not being a partner in the program, has a widely recognized history of success in educational management and also appears in the top placings. This evidence adds strength to the argument that management makes a difference to the advancement of education.

for school leaders and their staff.[43] There is still limited experimental evidence about the effects of effective instructional leadership in Brazil. Therefore, more work must be done to differentiate the effect effects of practices directly linked to the work with the teachers in the classroom (the instructional leadership) from the more general management and leadership practices, that affect the entire school space.

This impact evaluation helps us understand the power of effective instructional leadership. For this purpose, Jovem de Futuro developed a "pedagogical package"[44]. All schools will receive this "pachage", but a random draw was made to define some that will receive first than others, creating also a treatment and comparison group to understand the effects of this intervention.

43 Some of the most important studies that corroborate this conclusion are: Fullan, 2001; Leithwood et al., 2004; and Robinson, Lloyd and Rowe, 2008.

44 The pedagogical package is made up of 7 actions: (1) pedagogical feedback integrated into the '*Foco Brasil*' ('Brazil Focus') platform, (2) new training for the pedagogical coordinator, (3) a protocol directed at the schools designed to prevent students dropping out of school, (4) a protocol directed at the schools designed to welcome the students into the 1st year (those who reach the 9th year with gaps in their learning), (5) a protocol directed at the schools for the effective use of the collective planning period with the teachers, (6) practically-focused communities to involve school principals, pedagogical coordinators, supervisors and young people, which will address pedagogical matters, and (7) a coexistence program in the school.

# CONCLUSION

The evaluation of the Jovem de Futuro shows that the program has improved achievement and graduation rates in secondary school and has reduced inequality. These results are especially important considering the scale of the program and the different contexts in which it has operated. The impacts identified are similar to those included in international studies that have also investigated the effect of school leadership and management on students' performance.

Despite all the efforts made to use the most scientifically rigorous impact evaluation resources available, the Instituto Unibanco is aware that no method is perfect, nor can any method provide all the answers necessary to make the program perfect. This is especially true of something as complex as student learning. Investments in different evaluations have been made over the last five years, with results that will be explained in a future publication.

However, the experimental evaluation, when rigorously applied, plays an important role in providing the best evidence possible of the impact of a program on the targeted public. It is also essential in providing support for an ongoing reflection on the functioning of the program itself, in the pursuit of continual improvement.

For the Instituto Unibanco, continuous improvement requires evidence-based, data-driven decisions. This belief is the foundation of the Jovem de Futuro program. By transparently publishing the results of this evaluation, Instituto Unibanco is also looking to fulfill its mission to be accountable to society in relation to the result of the efforts made in partnership with the states, and thereby contributing to the public debate with respect to its most important goal: to guarantee the right to learning to every young person in Brazil.

# BIBLIOGRAPHY


ANGRIST, J. D.; PISCHKE, J. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press, 2008.

_____ et al. Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment. *American Economic Review*, Nashville, v. 92, n. 5, p. 1535-58, 2002.

_____; BETTINGER, E.; KREMER, M. Long-term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia. *American Economic Review*, Nashville, v. 96, n. 3, p. 847-62, 2006.

ÄRLESTIG, H.; DAY, C.; JOHANSSON, O. (Eds.). *A Decade of Research on School Principals: Cases from 24 Countries*. New York: Springer, 2016.

BANERJEE, A. et al. Can Information Campaigns Raise Awareness and Local Participation In Primary Education? *Economic and Political Weekly*, Mumbai, v. 42, n. 15, p. 1365-72, 2007.

BARROS, R. P. Desenho da avaliação de impacto do Projeto Jovem de Futuro. São Paulo: Instituto Unibanco, 2016.

_____ et al. O impacto do Projeto Jovem de Futuro sobre a aprendizagem em escolas públicas. São Paulo: Instituto Unibanco, 2016.

_____ et al. Assessment of the impact of the Jovem de Futuro program on learning. New York: The World Bank, 2018.

_____ et al. Avaliação de impacto da iniciativa Jovem de Futuro sobre o aprendizado. São Paulo: Instituto Unibanco, 2020 (in print).

_____; FRANCO, S. *Relatório metodológico: processo de aleatorização da avaliação de impacto do Jovem de Futuro*. São Paulo: Instituto Unibanco, 2018. Available at: <https://cedoc.observatoriodeeducacao.org.br/item/?cod=123456789_5232>. Accessed on: May 8, 2020.

BLOOM, N. et al. Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, Cambridge, v. 128, n. 1, p. 1-51, 2013.

BLOOM, N.; LEMOS, R.; SADUN, R.; VAN REENEN, J. Does Management Matter in Schools? *The Economic Journal*, London, v. 125, n. 584, p. 647-74, 2015.

BLUNDELL, R.; BOZIO, A.; LAROQUE, G. Extensive and Intensive Margins of Labour Supply: Work and Working Hours in the US, the UK and France. *Fiscal Studies*, London, v. 34, n. 1, p. 1-29, 2013.

BOTHWELL, L.; GREENE, J.; PODOLSKY, S.; JONES, David. Assessing the Gold Standard: Lessons from the History of RCTs. *New England Journal of Medicine*, Boston, n. 374. p. 2175-81, 2016. DOI: <10.1056/NEJMms1604593>.

BRASIL. INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Matrizes e escalas. Available at: <http://provabrasil.inep.gov.br/escalas-de-proficiencia>. Accessed on: May 2, 2020.

_____. Sistema de Avaliação da Educação Básica (Saeb), 2017. Available at: <http://portal.inep.gov.br/educacao-basica/saeb/resultados>. Accessed on: May 2, 2020.

BRASIL. MINISTRY OF EDUCATION (MEC). National Education Plan, 2014. Available at: <http://portal.mec.gov.br/arquivos/pdf/pne.pdf>. Accessed on: May 2, 2020.

COLEMAN, J. S. et al. *Equality* of Educational Opportunity. Washington: US Government Printing Office, 1966.

DEMING, W. Edwards. **Out of the Crisis.** MIT Press, 2018.

DOBBIE, W.; FRYER, R. G. Getting Beneath the Veil of Effective Schools: Evidence from New York City. *The American Economic Journal: Applied Economics*, Santa Clara, v. 5, n. 4, p. 28-60, 2013.

DUFLO, E.; DUPAS, P.; KREMER, M. School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools. *Journal of Public Economics*, Amsterdam, v. 123, p. 92-110, 2015.

_____. Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, Nashville, v. 101, n. 5, p. 1739-74, 2011.

DUNCAN, G. J.; MAGNUSON, K. Investing in Preschool Programs. *The Journal of Economic Perspectives*, Nashville, v. 27, n. 2, p. 109-32, 2013.

FLESSA, J.; BRAMWELL, D.; FERNÁNDEZ, M.; WEINSTEIN, J. School Leadership in Latin America 2000-2016. *Educational Management Administration & Leadership*, London, v. 46, n. 2, p. 182-206, 2018.

FULLAN, M. Large-Scale Reform Comes of Age. *Journal of Educational Change*, Dordrecht, n. 10, p. 101-13, 2009.

FRYER, R. G. Injecting Charter School Best Practices in to Traditional Public Schools: Evidence from Field Experiments. *The Quarterly Journal of Economics*, Cambridge, v. 129, n. 3, p. 1355-1407, 2014.

_____. Management and Student Achievement: Evidence from a Randomized Field Experiment. *NBER Working Papers*, Cambridge, n. 23437, 2017. Available at: <https://www.nber.org/papers/w23437>. Accessed on May 2, 2019.

GERTLER, P. J. et al. *Impact Evaluation in Practice*. New York: The World Bank, 2016.

HENRIQUES, R. M. S.; ROCHA, H. C. Ensino médio no Espírito Santo: juventudes e garantia do direito à aprendizagem. In: HARTUNG, P. (org.). *Espírito Santo: como o governo capixaba enfrentou a crise, reconquistou o equilíbrio fiscal e inovou em políticas sociais*. Vitória: Government of Espírito Santo, 2018.

HERMAN, R. et al. *Turning Around Chronically Low-Performing Schools: A Practice Guide* (NCEE 2008-4020). Washington: National Center for Education Evaluation and Regional Assistance; Institute of Education Sciences, 2008.

HILL, P. J. et al. Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives*, Chicago, v. 2, n. 3, p. 172-77, 2008.

INSTITUTO UNIBANCO. *Jovem de Futuro* Project, 2010.

_____. *Ensino Médio Inovador e Jovem de Futuro: por uma escola mais participativa, eficiente, criativa e transformadora*, 2013.

_____. *O Plano em Ação: circuito de gestão contínuo, integrado e com foco nos resultados de aprendizagem*, 2013.

_____. Sumário executivo sobre a memória e história do Projeto Jovem de Futuro, 2015.

_____. *Investimento Social em Larga Escala*, 2020 (In print).

_____. *Teoria da Mudança da Gestão em Educação*, 2020 b (In print).

LEITHWOOD, K. et al. *How Leadership Influences Student Learning*. Minneapolis; Toronto: Center for Applied Research and Educational Improvement; Ontario Institute for Studies in Education, 2004.

TODOS PELA EDUCAÇÃO. Cenários da educação. Available at: <https://www.todospelaeducacao.org.br/pag/cenarios-da-educacao>. Accessed on: May 2, 2020.

_____. *De olho nas metas: primeiro relatório de acompanhamento das 5 metas do movimento todos pela educação*. Organized by the Expert Commission of the '*Todos pela Educação*' agreement. São Paulo, 2008.

ROBINSON, V. *Student-Centered Leadership*. San Francisco: John Wiley & Sons, 2011.

ROBINSON, V. M. J.; LLOYD, C. A.; ROWE, K. J. The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects of Leadership Types. *Educational Administration Quarterly*, Columbus, v. 44, n. 5, p. 635-74, 2008.

WEINSTEIN, J.; HERNÁNDEZ, M. Birth Pains: Emerging School Leadership Policies in Eight School Systems of Latin America. *International Journal of Leadership in Education: American Economic Review*, London, v. 19, n. 3, p. 241-63, 2016.


9 786586 069532