

INSTITUTO
ARRECIFE



PUBLIC POLICY UNDER DEBATE

POP

#01

NOVEMBER
2025

QUALITY OF EDUCATION, SAEB AND IDEB:

**the twists and turns of the indicator that became the
thermometer of Brazilian schools.**

Luiz Carlos Faria da Silva



Quality of Education, SAEB and Ideb:

the twists and turns of the indicator that became the thermometer of Brazilian schools.

Luiz Carlos Faria da Silva*

* A native of Rio de Janeiro (Carioca), holds a degree in Pedagogy from the Pontifical Catholic University of São Paulo (1985), a master's degree in Education: History, Politics, Society from the Pontifical Catholic University of São Paulo, PUC-SP (1993), and a doctorate in Education from the State University of Campinas, UNICAMP (2000). He is an Associate Professor at the State University of Maringá, UEM. He has experience in the field of Education, with an emphasis on Philosophy of Education and Literacy Teaching Methods.

EDITORIAL BOARD

Alexandre Marques
Alexandre Vasconcelos
Eduardo de Alencar
Luiz Ramiro
Sávio Pontes

EDITOR

Eduardo de Alencar

PROOFREADING

Antunes Fernandes

FREQUENCY

Four-monthly

EDITION

POP - Public Policy, No. 1. November 2025.

INSTITUTE SERIES

Series 1 - CPD - Police Sciences in Debate, No. 1

Series 2 - SN - National Sovereignty, No. 1

Series 3 - CPD - Police Sciences in Debate, No. 2

Series 4 - POP - Public Policy, No. 1

Original signed by the author in Maringá, July 26, 2025. The opinions expressed in this document are the responsibility of the author and do not necessarily reflect the position of Instituto Arrecife or its Editorial Board.

Contents

1.	Introduction	4
2.	SAEB: the Basic Education Assessment System	7
3.	The Ideb: a project whose execution reveals design flaws?	9
4.	Notes	14

1. Introduction

In 2007, with the aim of improving the quality of school education in Brazil, the Ministry of Education – MEC – launched a set of education policies. In that context, **INEP – the National Institute for Educational Studies and Research Anísio Teixeira** – introduced the *Ideb – Basic Education Development Index*.¹ The original intent of the Ideb was to give society and education policymakers information about the evolution of education quality. Its calculation had, in summary, the following mathematical expression:

$$\text{Ideb } ji = N_j \times P_j, 0 \leq N_j \leq 10 \text{ and } 0 \leq P_j \leq 1.$$

Mathematical formulas frighten many people, but that is not always warranted. This formula is a case in point. The Ideb of unit j in year i equals the product – that is, the multiplication – of the numerical value of N , for unit j , at the end of year i , by the numerical value of P , for the same unit and at the end of the same year, where N is greater than or equal to 0 and less than or equal to 10, and P is greater than or equal to 0 and less than or equal to 1. A simple multiplication.

The numerical value of the capital letter **N** corresponds to the **simple average** of the proficiency² scores on the

SAEB – the Basic Education Assessment System – at the end of each stage of this first of the two levels into which national education is organized. N expresses student performance: it looks at the school, focused on student learning. The numerical value of the capital letter P , meanwhile, without losing sight of the student, looks at the efficiency³ of learning production, concerned with the time spent generating it. P expresses the average pass rate of students, taking into account the grades within each teaching stage – that is, the movement of enrollments between school grades: school flow. N is an arithmetic mean that expresses learning at school, and P is a rate that expresses the quality of the use of resources and time to generate learning. We learn to calculate arithmetic means and rates in elementary school.

The lowercase letter j indicates the education unit⁴ for which the Ideb is calculated. The lowercase letter i refers to the year in which SAEB was administered at each education unit.⁵ In the formula, N and P represent the main pieces of information. j and i , written as lowercase letters set below the imaginary baseline, represent secondary information.

It is clear, then, that the numerical value of the Ideb was meant to indicate the quality level of school education. That level would result from the product of numbers expressing two aspects⁶ of school educational action: the cognitive value added, and the optimization of the use of means to achieve that addition. Given this, we should understand two things: first, how the Ideb became the thermometer for the evolution of education quality; second, how it was concluded that the desired quality level would be reached once that thermometer read **6.0**.

Implicit in the Ideb formula is the following idea: in a perfect school, everyone would develop 100% of the competencies and skills that the teaching program, designed on the basis of the school curriculum, aimed at and established. No student would need more than one school year to successfully complete a given grade's course of study. A maximally defective school, by contrast, would be one in which **everyone** took more than one school year to get through a grade and **no one** managed to complete the course of study successfully. These would be two extreme conditions, unlikely to be found in reality. In any case, any factor of 0 in the Ideb formula – whether in the performance measure or in the flow measure – would zero out the product and yield an Ideb of 0. Conversely, values of 10 for performance and 1 for school flow would maximize the product and yield an Ideb of 10.

In the same year it launched the Ideb, MEC launched three further initiatives: Educacenso, the Compromisso Todos pela Educação [All for Education Commitment] Plan of Goals, and the PDE – Education Development Plan. The first was meant to bring precision and reliability to school statistics; the second, to unify and drive convergence among inter-federative actions to improve education; and the third, to publicize the systematization of these actions and reinforce the perception that they formed an organic, interwoven, and complementary set of education policies. Tying these initiatives together, one decision was built into the Compromisso Todos pela Educação Plan of Goals, laid down in Decree 6.094/2007: the Ideb would be the official indicator for gauging the evolution of education quality.

To use the Ideb as a “thermometer” of education quality, INEP took the following steps:

- 1- For each teaching stage, since the data were available from SAEB and the School Census, the mean between the numerical values of SAEB performance (Portuguese Language and Mathematics) and the numerical values of pass rates was calculated, yielding an estimated average Ideb for Brazilian education in 2005, referred to here as t2005. The values were 3.8, 3.5, and 3.4, respectively, for the completion of the initial years of elementary school, the final years of elementary school, and high school. 7
- 2- The average quality of school education in OECD – Organisation for Economic Co-operation and Development – countries was taken as the benchmark.
- 3- Two target figures would express this benchmark: the average proficiency of students in OECD – Organisation for Economic Co-operation and Development – countries on the PISA – Programme for International Student Assessment – Reading and Mathematics tests, and the pass rate in basic education in those countries.
- 4- The Reading and Mathematics proficiencies on SAEB tests equivalent to the PISA target proficiencies were estimated. Their numerical values, expressed on a scale of 0 to 500, were converted, through mathematical calculations, to a scale of 0 to 10.8 Next, the arithmetic mean between them was calculated, and the OECD countries’ pass rate was taken: 0.96. These would be the values of N and P entering the Ideb formula.

- 5- Applying these values – the arithmetic mean of the Portuguese Language and Mathematics proficiencies expressed on the 0-to-10 scale, and the pass rate expressed as the ratio between the number of students passing and total enrollment – the value 6.0 was obtained. The Ideb thermometer should read this value in t2021, t2025, and t2028, respectively for the initial years of elementary school, the final years of elementary school, and the completion of high school.
- 6- INEP set t2017 as the start of the count toward reaching the target, and staged the necessary effort into two broad time waves. In the first period, t2007 to t2016, education policies would be aligned. In the second, the objectives would be reached in three cycles, beginning with students enrolled in the 1st year of elementary school in 2017:

Initial years of elementary school: reach Ideb 6.0 in 2021.

Final years of elementary school: reach Ideb 6.0 in 2025.

High school: Ideb 6.0 in 2028.9

Figure 1
Timeline of Ideb targets and when the initial-years target was actually achieved



Source: prepared from the INEP intermediate targets described in this article and the SAEB 2023 results.

Seven years later, Law 13.005/2014 renewed the PNE – National Education Plan –, in force from 2014 to 2024. The law mandated continuous evaluation of the execution of the plan’s targets and of the strategies used to reach them. It stripped the Ideb of its prerogative as the exclusive indicator of the evolution of education quality¹⁰ and mandated periodic monitoring and evaluation of the plan’s execution and the targets’ attainment.

By the end of 2016, when it was published, the Report on the 1st PNE Monitoring Cycle, covering the 2014–2016 period, warned, on its page 186:

“it is important to pay attention to Strategies 7.1 and 7.2 of Target 7 of the PNE, in order to establish the common national curriculum base and define the sufficient and desirable levels of learning, in relation to the learning rights and objectives and the development expected for each year of study. Based on these actions, it will fall to education systems to make efforts to guarantee every student the right to learn.”

In 2017, for Early Childhood and Elementary Education, and in 2018, for High School, the BNCC – National Common Curricular Base – was approved. Learning rights and objectives were legally defined. The Report on the 2nd PNE Monitoring Cycle, published in 2018, turned the volume knob clockwise on the pressure to define the learning level, so that an accurate assessment could be made of the attainment of Target 7 and the application of two of its strategies (page 148):

“In this context, it is worth noting the achievement of Strategy 7.1 of Target 7, with the enactment of the National Common

added) to define the ‘sufficient’ level of learning in relation to the learning rights and objectives and the development expected for each year of study, as called for by Strategy 7.2. Only in this way will it actually (emphasis added) be possible to monitor the learning of basic-education students and guarantee everyone the right to learning.”

This pressure was repeated in the following reports, from 2020 and 2022, but it was no longer as severe in the 2024 report:

Although the fact of (sic!) the absence of an official definition of the “sufficient” level of learning limits full monitoring of Target 7, (emphasis added) the data presented here and the evidence already produced by several studies in the field indicate that the country is still far from ensuring that every elementary- and high-school student reaches a sufficient level of learning in relation to the learning rights and objectives and the development expected for their year of study, as called for by Strategy 7.2. (page 187).

Invariably, these reports underscored two points. The first is that the Ideb was rising, but learning was not. The second is that INEP resisted indicating the minimum learning levels required at each teaching stage in terms of SAEB proficiency. Without that, it was not possible to gauge the effectiveness of the use of Strategy 7.2 of Target 7 of the PNE.¹¹

But years after the start of the ambitious plan to raise the quality of national school education, once 2020 arrived, the Union’s contribution to FUNDEB increased and a new system for distributing these resources was established. This heightened the urgency of an evaluation of how the Ideb was being used.

Made up of resources drawn from part of the tax revenues of States and Municipalities and supplemented by resources from the Union’s tax revenue, FUNDEB tops up resources whose overall volume of application in Education is mandated by the constitutional text and regulated by law.¹²

Under the new FUNDEB law, the increase in the Union’s contribution will take into account the quality and equity of the education provided by the federative entities.¹³

Specifying the conditions to be met in order to access these resources gave rise to the creation of new indicators of the development of education quality, and – as the PNE action-monitoring reports do – insisted on knowing the proportions of students with sufficient learning – in the technical jargon of large-scale assessments, standard setting.¹⁴

We are, then, at a propitious moment to assess at least two points regarding the Ideb. First, we can ask whether it has proven to be an adequate and useful indicator of the development of education quality. Second, we can reflect: why does the Ideb rise, when its calculation formula uses average SAEB proficiencies, if those average proficiencies are falling or holding at the poor level they were at in 1995, 12 years before it was established?¹⁵

Since the use of the Ideb is embedded in education policies that have been in place for decades, in which the right to education and the guarantee of that education’s quality are intertwined,¹⁶ and since one cannot discuss the Ideb without reference to SAEB, in order to better answer the questions raised in the previous paragraph we will take two steps. In the first, since Brazil has practically left behind the problem of adult illiteracy and has given practically its entire child and adolescent population access to school,¹⁷ and has since turned to stare fixedly at the problem of the quality of the school education provided, we will see how recent education-policy decisions, in the very field of large-scale assessment, neutralized the potential use of SAEB data to obtain a view of that quality that is both synchronic and diachronic, and squandered SAEB’s own capacity to support the planning and execution of effective educational interventions. In the second step, we will focus directly and specifically on the Ideb to show why it has not proven to be a good indicator of the evolution of school education quality, nor is it fulfilling the role for which it was designed.

SAEB: the Basic Education Assessment System

At the start of the 1980s, Brazil still faced the problem of unequal access to schooling, even as that access was becoming more widespread, since the massive influx of children into school was much more intense in the urban centers of the more industrialized regions. Amid concern over school dropout between the first and second years of what is now elementary school,¹⁸ then called “primeiro grau” education, there was great concern over the provision of educational inputs and over the inequalities in educational provision, both regional and between social strata. These inputs were, above all, school buildings and facilities, libraries and laboratories, canteens and cafeterias, staff and administrative resources, curricula and teaching programs, qualified and credentialed teachers, as well as food, school transportation, and teaching materials.

In the course of a set of actions attempting to resolve these problems, it became urgent to know the average learning level produced by school units and education networks. It was necessary to measure the impact of ongoing interventions in the area of input provision – for example, in EDURURAL, a program to expand and improve education in rural Northeast Brazil. This foreshadowed a shift in perspective in how the quality of school education in Brazil was assessed: looking at outputs, that is, teaching results, rather than only at the increase in inputs, that is, educational inputs. And doing so by means of a direct indicator of outputs: tests that would measure the cognitive value added. That was the embryo of the idea of a SAEB in Brazil.¹⁹

After work carried out in various states, involving staff from the State Departments of Education, from MEC – the Ministry of Education – and from international agencies, SAEB took its first steps between 1990 and 1993. Under INEP’s leadership, tests were designed and a set of information to be gathered through questionnaires applied alongside the tests was defined. Cross-referencing these two sets of information would provide a clearer picture of the impact of each factor on learning.

After a few rounds, INEP validated and consolidated both the psychometric framework of the tests applied in SAEB – which uses IRT – Item Response Theory – mathematical models to analyze the data generated, an approach applied from 1995 onward – and the pedagogical approach to teaching and learning, adopted in SAEB tests

PCNs [National Curriculum Parameters], a decision applied from 2001 onward.²⁰ We note, however, the existence – likewise validated and consolidated – of pedagogical and psychometric approaches distinct from those adopted in SAEB. And we underscore how unknown and disregarded they are in the national debate on the quality level of basic education in the country.²¹

If the cognitivist-sociointeractionist pedagogical approach to teaching/learning, on the one hand, holds that the process of cognition involves the informational processing of data from the environment and constant adaptations and re-adaptations – both internal, associated with biopsychic maturation, and external, associated with encountering the social environment, especially in what is called the “social practice of language use” – the psychometric approach of IRT, on the other hand, assumes that external, directly observable behaviors reflect latent traits of internal characteristics that are observable only indirectly. In this case, the response to a test item, an observable behavior, would be associated with the latent trait of an internal characteristic or capacity, observable only indirectly.²²

In analyzing response behaviors to test items, IRT starts from the general premise adopted in latent trait theory: the probability of a correct response rises or falls as the magnitude of the latent trait required to choose the correct answer rises or falls. That magnitude can theoretically range between immense and minuscule. This is the case, for example, with reading and text comprehension. There are beginning readers, skilled readers, and outstanding readers or experts. The probability of correctly answering a text-interpretation item would theoretically range from close to 0, for the first type of reader, to close to 1, for the last. A probability of 0 would mean total impossibility of a correct answer; 1, a guaranteed correct answer.

In practice, these extreme conditions are implausible. Just as, in the pedagogical domain, there would be no magnitude of null or total learning performance expressed as the minimum or maximum numerical value on the SAEB proficiency scale, marked on a ruler from 0 to 500 points, likewise, in the psychometric domain, there would be no magnitude of latent trait entirely associated with the probabilities of

nonexistent or complete success in answering a test item.²³

This means there are two perspectives involved in the application of tests. The perspective of Psychometrics, interested in gauging the magnitudes or degrees of knowledge, skills, and attitudes that, according to external observation, vary along a continuum; and the perspective of Education, interested in understanding the effects of pedagogical/didactic intervention on increasing the continua of knowledge, skills, and attitudes, likewise observable in teaching's susceptibility to increase those magnitudes or degrees of knowledge.²⁴

Properly understood, then, in SAEB the perspectives of the psychometric approach and the pedagogical approach converge.

The use of computational and information-technology means for scoring and processing information collected from tests administered to large populations makes it possible to process, simultaneously, a much larger quantity of data than any individual teacher could analyze on their own within their classes. By coding information on each student's correct and incorrect answers and interpreting the data with the help of the mathematical models provided by IRT, large-scale assessments are able to identify and describe, with reasonable precision, the average cognitive profile added, at each teaching stage, for the school populations of interest.²⁵ It is on the basis of this information that the effectiveness and efficiency of education policies can be judged. SAEB has the potential to support informed judgment about education policies aimed at raising average cognitive profiles across entire school populations.

The analyses of SAEB test results, presented in reports published by INEP, favor observing the numerical values of the average proficiencies obtained in each cognitive domain at the end of each teaching stage. For those responsible for education policy, it is important to know these average proficiencies. For the internal public of education units, it is important to know, besides the average proficiencies, the proportions of students whose proficiencies fall within each of the levels into which the scale is pedagogically interpreted.²⁶ For the general public, meanwhile, it is crucial to know the proportion of proficiencies above and below a minimum expected threshold for each teaching stage, at each education unit, and how that proportion changes with each test round and over the long term.

Based on the analysis of the SAEB 2001 results, INEP decided to refine the data analysis and go beyond simply publishing average proficiencies – a decision later consolidated in the SAEB 2003 National Report, published in 2006. It defined five labels for learning performance according to the thresholds reached by proficiencies, identifying stages of knowledge acquisition:

very critical, critical, intermediate, adequate, and advanced.²⁷ Three special reports, published between 2003 and 2005, inaugurated this way of reporting results and explicitly reported the percentage of students performing at each stage. In the first of these documents, INEP recorded:

*"The disclosure of Saeb results emphasized the performance average, but did not classify that average into stages of knowledge acquisition, as is done below. Without indicating the stage of knowledge acquisition in which the average falls, it conveys little information. The analysis of the proficiency results drawn from the tests applied in Saeb 2001 was the basis for building five performance categories: very critical, critical, intermediate, adequate, and advanced."*²⁸

The public could thus see that, with the exception of what occurred in the 8th grade in Portuguese Language, more than half of Brazilian students were at the very critical and critical stages of knowledge acquisition.²⁹

Quickly, however, INEP abandoned this approach. The SAEB 2005 report, published in 2007, went back to centering the analysis of test results solely on comparing average proficiencies.³⁰

The information produced by SAEB would have greater potential to guide the effort to raise the quality of education in the country and more strongly engage everyone interested in pursuing better results, first, if INEP took on the responsibility of officially publishing the minimum proficiencies expected of students completing each stage of schooling; second, if it disclosed the percentages of students performing above and below those cutoff lines.³¹ To do so, it would have to reject the temptation to frame the reporting of results within an analytical matrix that sociologically justifies poor performance.³²

It so happens that, instead of making good use of SAEB data, INEP instituted and adopted the Ideb – Basic Education Development Index.

We will see how this initiative created a great uproar, squandered efforts already under way within SAEB and, as reality confirms today, is failing. Brazil reached, even if late, in 2023, rather than in 2021 as expected, the self-set target of an average Ideb of 6.0 for the initial years of elementary school. Even so, average performance in Reading and Mathematics is dismal. And the claimed rise in education quality in the initial stage of elementary school does not carry over to the following stage, the completion of elementary school. Even less so to the completion of basic education, as is well known.³³

The Ideb: a project whose execution reveals design flaws?

The Ideb became, as we have seen, the thermometer for the advance of national education quality along the path of implementing policies stitched together in the PDE and the Compromisso Todos pela Educação, launched in the same year, at a time when the use of evidence was spreading across the country more as a fad than as a culture. It was born with a brilliant future already contracted.

The initiative had strong appeal and support because it not only seemed to offer a compass for reaching a highly valued goal – namely, a high quality level for national basic education – but, apparently realistically, proposed to reach that goal within a timeframe suited to the scale of the task: a decade of policy alignment before the count began, and, from that point, the time needed for each cohort to move through the corresponding teaching stage. There would be ample time to prepare, since the measuring clock would only start running in 2017, as we noted in the Introduction. With the added “luxury” of the always-expected benefits from the rise in preschool enrollment rates, since school attendance among 4- and 5-year-olds accelerated: between 2005 and 2015, it went from 52.4% to 80.0%, in the annual PNAD series used by IBGE’s Synthesis of Social Indicators, distinct from the continuous PNAD series cited in notes 17 and 52. This growth, intense among the poorest, raised expectations of a positive impact on later teaching stages, since:

“there was a decrease in the inequality of access to school among per capita monthly household income quintiles. The difference between the proportion of 4- and 5-year-old children attending school in the highest-income quintile and in the lowest-income quintile fell by half – that is, it went from 34.3 percentage points to 14.0 percentage points between 2005 and 2015.”³⁴

The execution of the PDE, in which the use of the Ideb featured as the star, was planned to be gradual, that is, calibrated over time. The overall objective would be reached in three cycles, the last in 2028, when the students enrolled in the 1st year of elementary school in 2017 would reach the 3rd year of high school.³⁵ Expectations were very positive because, after all, the Ideb was an easy-to-understand indicator that could rally and galvanize educators’ efforts and mobilize society’s support.

Favorable public opinion then formed. First, around the idea of setting an overall target for the development of education; second, around the strategy of several steps to be climbed in that development; third, around the idea of using a tool capable of gauging the pace of progress in education quality and calibrating the effort needed to reach the target. All of this was stitched together by another idea: the commitment of education stakeholders to results, called accountability among specialists. The stage was set. The Ideb’s general concept stemmed from the pressure that education policymakers began to face in an era when fiscal balance had become a leitmotif. Everyone was calling for adequate learning within a reasonable time, since that would amount to optimizing the use of public resources.

As early as 2007, the PDE stated that an “automatic-promotion industry” had formed in Brazil.³⁶ This would distort real knowledge of school flow. Educacenso and the new School Census resolved the problem.³⁷ Brazil gained better information conditions regarding the real state of enrollment numbers and their movement between grades across teaching stages. The same happened with the consolidation of SAEB. It had already given Brazil a tool for knowing the average cognitive value added at each teaching stage of school education.

By adding the Ideb to these initiatives, MEC and INEP accelerated the development of a culture of evidence use and accountability. There would now be a single indicator in which information on these two dimensions of school education quality – performance and school flow – would come together. Those responsible for education policy, the promoters and disseminators of these policies in so-called civil society, advocacy systems, and the media all fell into line. The initial “success” of the PDE’s proposals was mirrored in the evolution of the Ideb and crossed borders.³⁸

But before we look at what happened over the medium term, from the standpoint of the Ideb’s capacity to monitor and guide the improvement of school education quality in the country, it should be noted that:

The most important point in the formulation initiative – in the literal sense – of the Ideb was the choice of education quality in OECD countries as the touchstone for education quality. And the use of PISA as a mirror of that quality. It would be necessary to seek in SAEB a

The procedure for finding this PISA target score is described in a Methodological Note on making PISA performance compatible with the SAEB scale.³⁹

INEP's technical note on making PISA and SAEB performances compatible for purposes of setting the Ideb's target value, apparently drafted in-house at INEP, does not indicate either an author or the INEP unit responsible for preparing it.

In the first two paragraphs, the note states what it calls the compatibility between PISA and SAEB performance levels and describes the procedure for achieving it.

In the third and fourth paragraphs, it acknowledges the existence of limitations in carrying out this procedure.

In the fifth paragraph, however, it announces that it will disregard these limitations.

Having chosen the PISA target scores, Reading and Mathematics, the corresponding SAEB "scores" would be sought for the 8th grade of elementary school, which at the time lasted eight years. The figure corresponding to the pass rate in OECD countries, 0.96, had already been chosen. With these figures in hand, it would be possible to arrive at the SAEB performance level that Brazilian education should have for its quality to be comparable to that of OECD countries.

a "score" that reflected a target score in PISA. For school flow, the reference was the average pass rate of OECD countries: 0.96. OECD countries' indicators were thus used to anchor Brazil's indicators.

Let us now look at some elements involved in the specific procedure adopted by INEP to make SAEB and PISA performances compatible with each other.

Like SAEB, PISA has scoring scales divided into levels for Reading and Mathematics. In SAEB 2003, the scales were pedagogically interpreted into 8 and 10 numerical scoring intervals, respectively for Portuguese Language and Mathematics.⁴⁰ Later, there were changes to these levels of pedagogical interpretation. With the adaptation of the SAEB Reference Matrices to the BNCC, there will be further changes starting with the SAEB 2025 round.⁴¹ In PISA 2003, the scoring values on the scales were interpreted into 5 and 6 numerical intervals, which pedagogically mean, respectively, 5 levels of competencies and skills in Reading and 6 in Mathematics. These levels are cumulative. Anyone who has developed the competencies/skills described at the higher levels also masters those described at the lower levels. Today, PISA pedagogically interprets skills, in both Reading and Mathematics, into 8 levels on the scale, 3 of them below level 2.⁴²

In PISA, since 2003, the proportion of Brazilian adolescents whose Reading performance was below the expected minimum remained stable, until 2022, at a disastrous level: around 50%. In Mathematics, with an insignificant decline between 2003 and 2022, this proportion likewise remained stable. But at a catastrophic level: from 75.2% in 2003 to 73.4% in 2022.⁴³

The average proficiency of Brazilian adolescents was 356 in Mathematics and 403 in Reading,⁴⁴ whereas, in OECD countries, it was 500 points in Mathematics and 494 points in Reading (OECD Average).⁴⁵ It was to the Reading and Mathematics average in PISA 2003 among OECD countries that INEP looked when it wanted to fix a reference point for equating with the SAEB 2003 proficiencies. But the numerical values of the PISA 2003 scores that INEP chose to equate with the SAEB 2003 scores, described as the "average obtained by participating" OECD countries in PISA, are not the ones that appear as the OECD Average in the PISA 2003 report, although the numerical difference is very small.⁴⁶

In any case, the PISA 2003 proficiency values that INEP took as a reference for seeking a parallel score in SAEB fall, as also happens with the official average Reading and Mathematics proficiencies of students from OECD countries, within level 3 of the PISA scale. But all these values hover around the floor of level 3 and the ceiling of level 2.

It happens that, also according to the same PISA 2003 report, numerical differences of this magnitude may not represent differences in terms of the range of competencies and skills developed. For example, with an average of 356 points in Mathematics in 2003, Brazil's performance cannot be considered different from the performance of students from Indonesia, with 360 points, nor from Tunisia, with 359 points: differences of 4 and 3 points, respectively. In Reading, with an average proficiency of 403 points, the performance of Brazilian students cannot be considered different from the performance of students from Thailand, with 420 points, nor from Serbia, with 412: differences of 17 and 9 points, respectively. This means that the difference between these countries' proficiencies, although it places them in different positions in the participants' ranking, is not large enough to indicate a difference in the range of cognitive competencies and skills developed.⁴⁷

Having chosen the PISA 2003 reference scores in Reading and Mathematics, INEP identified, in the distribution of proficiencies of students from OECD countries who took PISA 2003, the proportions of numerical proficiency values that fell above and below them. It then identified, in the distribution of the numerical values of the SAEB 2003 proficiencies, the proficiencies that produced the same result (proportions of numerical values below and above):

they would be equivalent. The conclusion was: “scoring an x” on SAEB would be equivalent to “scoring a Y” on PISA.

There are, however, limitations to estimating proficiency equivalence between different tests, such as PISA and SAEB. Although acknowledged in its own technical note, INEP disregarded them.

Let us look at some details.

PISA is what is called an age-based exam. Its intended target population is defined by age/schooling: between 15 years 3 months and 16 years 2 months at the time of the test, enrolled in some teaching program and/or school corresponding to at least the 7th year of basic education. Students from this target population took PISA 2003. By contrast, SAEB tests were administered to students enrolled in the 8th grade, the last grade of the then eight-year elementary school.

Let us not forget Brazil's high rate of age-grade distortion. Without age-grade distortion, Brazilian students aged 15 to 16 should be enrolled in the 1st or 2nd year of high school. But when one looks at the sample of the Brazilian population that took part in PISA 2003, one sees that it includes a smaller proportion of students with age-grade lag than actually exists in our population aged 15 to 16. This means the following: the Brazilian sample of students who took PISA 2003 had a better profile than the country's real school population in that age bracket. A PISA technical report indicates that Brazil had one of the largest contingents of population aged 15 to 16 enrolled in a grade below the one corresponding to the 7th year of study in basic-education programs.⁴⁸

In short, there are considerable differences between the populations subjected to SAEB and to PISA. Even so, in its technical note, INEP assumed that the populations were comparable. They were not.

Worse. Based on this hypothetical comparability between the proficiencies of 8th-grade students in SAEB and students aged 15 to 16 in PISA, INEP, in the same technical note, estimated PISA proficiencies for students in the 4th grade of elementary school and the 3rd year of high school in SAEB. This was for the purpose of calculating the Ideb's intermediate targets for the initial and final years of elementary school and the 3rd year of high school.

Another point. The constructs that SAEB and PISA seek to measure are clearly different. PISA test items aim to gauge the ease with which adolescents use their knowledge and skills to meet the challenges of everyday life, not their mastery of content taught in school curricula, as is the case with SAEB.⁴⁹ The emphasis is placed on

mobilizing resources, understanding concepts, and the capacity to act in various situations involving the cognitive domains assessed. The approach is, therefore, cognitivist. SAEB test items, even though Brazilian school curricula adopt the same cognitivist approach to teaching and learning, assess the competencies and skills associated with mastery of the study content described in school curricula.

At the end of this chain of estimates based on conjecture, INEP's technical note concludes that Brazil's basic education would have an average quality comparable to that of basic education in OECD countries once the numerical values of the average proficiencies of Brazilian students in SAEB, in Reading and Mathematics, converged toward the SAEB/PISA equivalence values estimated by the procedure to which the technical note alludes, without, incidentally, documenting the calculation procedures used.

The timeframes for reaching the targets are not, in fact, known to most stakeholders. And few know that the distribution of Ideb values among school networks and schools, across Brazil, is dramatically unequal. Thus, the time needed for all schools and networks to converge on the same Ideb of 6.0, across the three teaching stages, throughout Brazil – which would represent achieving a balance between quality and equity – was estimated at 91 years.⁵⁰

We are in 2025. 2028 is three years away. The most recent Ideb intermediate targets for the completion of elementary school and high school were set for 2021. For later years, since the targets were expected to be reached in 2025 and 2028, there should be intermediate targets. In any case, none of the targets set for the 9th year of elementary school and the 3rd year of high school has ever been reached. On this point, given the genuine public interest, a few points that tend to escape most people's notice need to be strongly underscored, since, at the very moment the attainment of the Ideb target for the initial years of elementary school is being celebrated, the real picture is, in general, as follows:

- The SAEB 2021 and 2023 rounds for the 2nd year of elementary school show that more than half of students are still not fully literate, most of them after four years of schooling, two of them in preschool;⁵¹
- Between 1995 and 2023, average SAEB proficiencies for the 5th year of elementary school rose 25.2 and 33.8 points, respectively in Portuguese Language and Mathematics – somewhere between half and two-thirds of the scale's standard deviation, a modest rise given the time elapsed and the volume of policies applied;
- The result of PIRLS 2021 Reading, which assesses students completing this same teaching stage, indicates

In the case of the two subsequent teaching stages, the 9th year of elementary school and the 3rd/4th year of high school, comparing SAEB 1995 with SAEB 2023, average proficiencies in Reading and Mathematics, respectively, either remained at the same poor level or are now lower than in 1995;

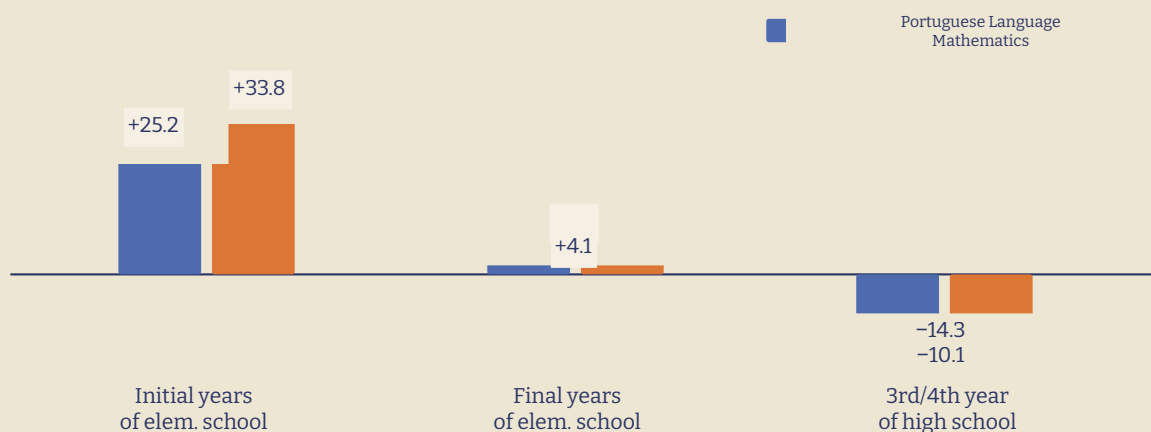
and in 65th place in Mathematics, with performance not distinguishable from that of countries ranked between 64th and 67th. And with 50% and 75% of proficiencies below the expected minimum level, respectively in Reading and Mathematics;

In the case of PISA, Brazil's average proficiencies place the country in 52nd place in Reading, with performance that is not distinguishable, in terms of competencies and skills, from that of countries ranked between 49th and 55th;

All of these figures are consistent with what INAF data show regarding literacy levels among the adult population.⁵²

Figure 2

Change in average SAEB proficiencies between 1995 and 2023, by teaching stage (scale points)



Source: prepared from the SAEB 1995 / SAEB 2023 comparison presented in note 15 of this article.

Beyond these points, which clearly show the Ideb's inadequacy for properly indicating the development of education quality, it is crucial to bear in mind that, in this scenario, the perception that the Ideb has failed in its purpose is only gradually and tacitly coming to be acknowledged. Proposals for improvements and course corrections that touch on the core of the problem are under way among those who worked on its advocacy.⁵³

On the official front, the identification of one of the effects of the limitations of its use, though not of its crucial flaw, appeared in analyses that INEP's own staff carried out as part of their duties monitoring the execution of PNE measures. The analyses pointed to the impossibility of verifying the use of Strategy 7.2 in fulfilling Target 7 of the PNE. The same flaw was underscored in technical notes authored by INEP staff dealing with the use of FUNDEB resources for financing education and with the use of the CAQ – Student Quality Cost – and the CAQi – Initial Student Quality Cost.⁵⁴ And it shows through, indirectly, in the back-and-forth seen in Administrative Rulings No. 556 of October 2, 2020, No. 828 of October 20, 2020, No. 119 of March 11, 2021, No. 918 of December 19, 2021, up to Administrative Ruling No. 188 of May 20, 2024, which set up an advisory commission to update the Ideb.

Constitutional Amendment No. 108/2020, Law 14.113/2020, and the amendment of some of the latter's provisions, made in items I and II of paragraph 3 of Law 14.276/2021, tie access to FUNDEB resources to SAEB performance, expressed in terms of the proportions of students with proficiencies above the minimum level expected for each teaching stage. But for this to take effect, INEP will have to

state the SAEB proficiency thresholds that will mark that minimum performance. However, the forces preventing INEP from doing so are the same ones that prevented recognition of the Ideb's crucial flaw: signaling progress in education quality at the very same time large-scale assessments were pointing to the persistence of dismal, even catastrophic, learning performance.

Although it is not limited to this, as we have just seen, at the outset the Ideb's error lay in the procedure used to link or equate proficiencies in SAEB and PISA. With its steps poorly documented, the undertaking amounted to a textbook case of violating the psychometric and statistical rules that safeguard the trustworthiness and reliability of this kind of disastrous Brazilian result in international comparison;

so challenging and sensitive an operation as estimating proficiencies from a test A across different versions of that same test A, or indeed on a test B, as was done in comparing proficiencies in SAEB and PISA.

This problem is not new. In the United States, the National Research Council, on behalf of the National Academies system,⁵⁵ published clear guidance for this type of operation, since there were several problems there related to equating scores across different tests. The issue is addressed in other well-established works.⁵⁶

Brazilian citizens need to free themselves from the web of influences through which the university education intelligentsia, the econometrists from financial institutions embedded in MEC, the bureaucratic-pedagogical establishment led by UNDIME – the National Union of Municipal Education Directors – and CONSED – the National Council of State Education Secretaries –, the NGOs of large economic-financial and media groups operating in education, and the school-material publishers – prisoners of the PNLD – National Textbook Program – tenders, tied to BNCC terms –, operating

interests that cross-fertilize one another, even if they do not always coincide entirely, have hijacked the nation's pedagogical awareness to the point of paralyzing it in the face of the barrage of negative information about the performance of our school education. This dynamic prevented recognition of the Ideb's underlying error, at the cost of nearly two decades of delay in the development of education. The network continues to operate, despite the growing difficulty of continuing to defend the importance attributed to this inept indicator.

Brazil needs to know the real state of school education in the country. And to learn to hold accountable those who, in the very context of introducing the concept of accountability into education in the country, led the nation into this mistake. It is urgent to put an end to the ethos of leniency toward evasions and attempts to redirect the arrow after it has already missed its target – which is what stakeholders in privileged positions and established policy makers are doing, all trying to “retread” the Ideb.

Professor Dr. Luiz Carlos Faria da Silva

Maringá, July 26, 2025

4. Notes

1. INEP is the federal agency responsible for producing and disseminating information, conducting studies, and disseminating statistics in the field of education. Reynaldo Fernandes was President of INEP between September 2005 and December 2009.
2. In school, each student has a certain level of aptitude — for example, in reading and written communication, or in using numbers and figures to solve mathematical problems. SAEB's tests detect and differentiate these various levels of aptitude and refer to them as proficiency levels. The word "proficiency" is used ambiguously in large-scale assessment reports.
3. See note 6.
4. We use the term educational unit to refer to both a school and a school network. The educational units for which INEP publishes the Ideb can be grouped by location, that is, Rural/Urban; by the territory's political-administrative organization, that is, State Capital/Interior; and by administrative dependency, that is, Public/Private.
5. SAEB administers tests every two years, in odd-numbered years. The results of these tests, expressed as average proficiencies, are comparable across editions.
6. These aspects are efficacy/effectiveness, on one hand, and efficiency, on the other. An intervention is considered efficacious/effective if it produces an effect in the expected direction. In the first case, under controlled conditions, for example, a laboratory or a classroom. In the second, under real conditions, such as a region or a school network. Efficiency, in turn, refers to the balance between the effort invested and the results produced. The subject is complex because, in many fields, especially in school education, it is not easy to clearly separate the contribution of the intervention from that of other, uncontrolled, intervening factors in producing an outcome. Two distinct interventions may both be considered efficacious/effective if they produce the effect in the expected direction, under controlled and real conditions. But between two effective interventions, one may be more efficient by producing a greater magnitude of effect, producing an effect across a larger number of circumstances and scenarios, or presenting a more favorable cost-benefit ratio. Cost, here, is not only financial: it is the relationship between the effort and resources employed and the benefit obtained. Cf. COCHRANE Brazil. Eficácia, efetividade e eficiência em saúde. In: *Estudantes para melhores evidências*. See <https://eme.cochrane.org/eficacia-efetividade-e-eficiencia-em-saude/> Note: Cochrane is a global, independent, and non-partisan network of health researchers, professionals, patients and caregivers who work to promote trust in high-quality health and healthcare information worldwide. It has a Brazilian branch. See <https://www.cochrane.org> and <https://brazil.cochrane.org/> Cf. Gottfredson, D.C., Cook, T.D., Gardner, F.E.M. et al. *Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation*. Prevention Science, Volume 16, Issue 7, Pages 893-926.
7. Cf. MEC/Inep. Ideb resultados. See at <https://ideb.inep.gov.br/>
8. Specific features of the mathematical formula used to convert SAEB proficiency scores from the 0-to-500 scale to the 0-to-10 scale tend to produce Mathematics proficiencies expressed as higher values than those for Portuguese Language. This may be leading to two distortions: an appearance of greater learning in Mathematics relative to Portuguese Language, which does not actually exist, and a ceiling effect on the averages. This is a technical issue that has already been raised, with a suggested corrective measure, in specialized publications. But INEP, as far as is known, has not acted on the issue raised or on the suggested corrective measure.
9. Everyone looks at the value of 6.0 for the Ideb, established as the target for Brazilian education. The target for the early grades was set, in INEP's trajectory of intermediate targets, for 2021; the year 2022, often repeated, comes from the PDE's rhetoric, which linked it to the bicentennial of Independence. Few, however, have examined the following points: the procedures used to arrive at the claim that an Ideb of 6.0 would represent good-quality education. There were 10 years of preparation to align policies before "starting the clock" on the race toward an Ideb of 6.0. This means: reaching the Ideb target for the 5th-year stage of elementary education in 2021 presupposed good-quality schooling since 2017, the year in which the students completing that stage in 2021 would have enrolled in its 1st year. The Ideb for the 5th-year stage of elementary education rose during the policy-alignment period, between 2007 and 2016, but did not reach the target in 2021. That year, it stood at 5.8. This figure, however, aggregates data from all public and private educational units in the country. Educational units run by private networks tend to show better indicators. Thus, publishing the Ideb by aggregating public and private networks may have produced a ceiling effect, that is, pushed the average upward. One should, however, be cautious with this claim, since there is no clarity about the extent of the aggregation. If it was total, that is, State Capital/Interior and Urban/Rural, in addition to Public/Private, we would need to recognize that educational units in the Interior and in rural regions tend to perform worse than those located in State Capitals and urban regions. This could potentially produce a floor effect. In any case, aggregating private-network data was not enough to reach the target in 2021. It was only achieved in 2023. See at <https://app.powerbi.com/view?r=eyJrIjoiaMGVjMzIwZWQzM2IzS00NmE0LTkwNjUtZjI1YjMyNTVhZGY0IiwidCI6IjI2Zjc2ODk3LWM4YWYWMtNGIxZS05NzhmLVVhNGMwNzc0MzRiZiJ9>
10. The main clause and sole paragraph of Article 3 of Decree 6.094/2007 had established that the Ideb would be the indicator used to gauge the quality of education. Paragraph 2 of Article 11 of Law 13.005/2014 qualified this standing of the Ideb.
11. National Education Plan 2014-2024. GOAL 7: Foster the quality of basic education at all stages and modalities, with improvements in school flow and in learning, so as to reach the following national averages for the Ideb. Strategies: 7.2) ensure that: a) by the fifth year after this Plan takes effect, at least 70% (seventy percent) of elementary and secondary school students have reached a sufficient level of learning in relation to the learning rights and objectives and the development appropriate to their year of study, and at least 50% (fifty percent) have reached the desirable level; b) by the final year this Plan is in effect, all elementary and secondary school students have reached a sufficient level of learning in relation to the learning rights and objectives and the development appropriate to their year of study, and at least 80% (eighty percent) have reached the desirable level. (Emphasis added by the author).

12. Constitutional Amendment No. 108/2020 (items a, b, and c of section V, Article 212-A) divided the Union's supplementary funding into three installments: the Fixed Annual Value per Student – VAAF – (item a), the Total Annual Value per Student – VAAT – (item b), and the Annual Value per Student by Result – VAAR – (item c). Only the last of these, the VAAR, is distributed according to learning outcomes and management conditionalities. These relate to the CAQ – Student Quality Cost – and CAQi – Initial Student Quality Cost – introduced into the constitutional text by that amendment, with the addition of §7 to Article 211. This is a difficult and complex subject – education financing. The 1988 constitutional text, in section VII of Article 206, introduced the guarantee of a quality standard as a principle on which teaching would be based, but specified nothing further. Law 9.394/1996, in section IX of Article 4 and in the main clause of Article 74, provided that specification. However, explicit references to the quality standard, embodied in the concepts of CAQ – Student Quality Cost – and CAQi – Initial Student Quality Cost –, only entered legislation through Strategy 6 of Goal 20 of the PNE – National Education Plan –, Law 13.005/2014. They were also referenced in Paragraph 2 of Article 48 of Law 14.113/2020, the new FUNDEB law. In the constitutional text, this specification only appeared with Constitutional Amendment No. 108/2020, which added §7 to Article 211. Even so, the use of the two parameters – CAQ and CAQi – is still not firmly established in the country's legal framework for education financing. This depends on regulation. Even today, in July 2025, the subject and the procedures for using these parameters are under debate in the Intergovernmental Commission on Education Financing, whose existence and responsibilities are set out in Articles 17 to 19 of Law 14.113/2020. All of the commission's actions can be followed at <https://www.gov.br/mec/pt-br/acao-a-informacao/participacao-social/conselhos-e-orgaos-colegiados/comissao-intergovernamental-fundeb>
13. Cf. Constitutional Amendment 108/2020, Law 14.113/2020, Law 13.005/2014, Decree 10.656/2021, and Ministerial Order No. 1/2024. Related to this subject is the calculation of an Initial Student Quality Cost – CAQi – and a Student Quality Cost – CAQ. These concepts were created within the National Campaign for the Right to Education, launched in 2002. The history of the debates and initiatives, including within the CNE – National Education Council – and the activity of the National Congress, along with a description of the current state of implementation of the CAQ and CAQi parameters, is set out in Technical Note No. 25/2020/CGIME/DIRED, from INEP. Available at https://www.gov.br/mec/pt-br/media/acao_informacao/pdf-arq/NotaTcnican252020CGIMEDIREDcomanexo1.pdf
14. Cf. Technical Note No. 12/2024/CGEF/DIRED-INEP: “SAEB has not yet produced this type of scale analysis for the stages considered under FUNDEB. However, since state-level assessments use the same SAEB scale for comparison purposes, the interpretations produced by these state assessments can be considered valid for reflecting SAEB learning levels.” What is being conceded here is the equivalence between performance on SAEB and performance on state assessments, such as SARESP. This kind of equivalence was strongly discouraged by a group of experts convened by the U.S. National Academies system when this same type of practice began there: establishing equivalences between performance on the NAEP – National Assessment of Educational Progress –, “their SAEB,” and performance on state tests. Cf. National Research Council. 1999. *Uncommon Measures: Equivalence and Linkage Among Educational Tests*. Washington, DC: The National Academies Press. The establishment of minimum performance standards at the end of each stage of schooling has been known, debated, and used for decades. Cf. Gregory Cizek, Editor. *Setting Performance Standards*. Foundations, Methods and Innovations. Second edition. Routledge, 2012. New York.
15. The decline, the stagnation at low levels, and the lack of steady gains across schooling stages emerge when comparing average SAEB proficiencies, 1995 versus 2023. The comparable historical series began in 1995. In the early-grades stage, 2023 proficiencies rose 25.2 and 33.8 points over 1995, in Portuguese and Mathematics respectively. In the final-grades stage, growth was near nil: 2.8 and 4.1 points, for Reading and Mathematics. At the end of basic education, 3rd/4th year of secondary school, there was a decline of 14.3 and 10.1 points, for Reading and Mathematics. This official, public data covers aggregated public- and private-network units. See charts, items 22-23, SAEB Press Kit 2019: https://download.inep.gov.br/educacao_basica/saeb/2019/presskit/PressKit_Saeb_2019.pdf. See also the SAEB Educational Panel: <https://app.powerbi.com/view?r=eyJrJoiOGZmMjNmOGQtOGFLYyO0Y2NhLWI4NmUtMGViZjAwNGJlNTAwIiwidCI6IjI2Zjc0ODk3LWM4YWMTNGIxZS05NzhmLWVhNGMwNzc0MzRiZiJ9> PIRLS 2021 International Results in Reading, Section 1: Countries' Reading Achievement. 57 countries; target population: 4th-year students – in Brazil, 5th year, by cohort average age. Over half of Brazilian students fall below the PIRLS midpoint (500); the Intermediate Benchmark is 475. Brazil averaged 419, ranking 43rd, though the report compared only 48 of 57 participants. Note: no trend for Brazil – our first participation. PIRLS runs in 5-year cycles; first edition 2001. Cf. IEA PIRLS 2021, at <https://pirls2021.org/results/achievement> PISA 2022 Results: The State of Learning and Equity in Education, Vol. I. In PISA 2022 (ages 15–16, at least 7th grade), Brazilian Reading proficiency, versus 80 countries, is as poor as or worse than in PIRLS: 52nd among participants, not significantly different from 49th–55th place. See https://www.oecd.org/en/publications/pisa-2022-results-volume-i_53f23881-en.html Despite this, the Ideb 6.0 target was met for 5th-year completers, two years late, in 2023; counting only state/municipal public units (80% of enrollment), the real Ideb fell further short. See <https://app.powerbi.com/view?r=eyJrJoiMGVjMzIwZWQzM2IzZS00NmE0LTkwNjUtZjI1YjMyNTVhZGY0IiwidCI6IjI2Zjc0ODk3LWM4YWMTNGIxZS05NzhmLWVhNGMwNzc0MzRiZiJ9>
16. The provisions that shape the political-institutional framework of the right to education, the procedures for developing education policies, and the linkage between these two processes are explicitly codified in the constitutional text, in Articles 211, 212, and 212-A, and in infra-constitutional legislation, for example, in Law 9.394/1996, Article 8.
17. The illiteracy rate among the population over age 15, in 2016 – a decade ago – stood at 6.7%. School enrollment rates by age, in the same year, were: 0 to 3 years, 30.3%; 4 and 5 years, 90.0%; 6 to 14 years, 99.2%; 15 to 17 years, 86.9%. Today, the illiteracy rate is 5.4%, and enrollment rates are, respectively: 38.7%, 92.9%, 99.4%, and 91.9%. Cf. IBGE. Continuous PNAD Panel. Education. Available at <https://painel.ibge.gov.br/pnad/c/>

18. At the turn from the 1970s to the 1980s, with school dropout between the 1st and 2nd grades rising to the number one spot among the problems of Brazilian schooling, the national university-based educational intelligentsia – which had already erred seriously during an earlier period of organizational effort to expand primary education across Brazil, between the 1930s and 1940s – erred again during the massification and universalization of access to what was then called primeiro grau (lower basic education). In fact, the bigger problem lay in grade repetition. Philip Fletcher and Claudio de Moura Castro. *Os alunos e a escola no Brasil de hoje*. Brasília, July 1986. Available for reading in the IPEA Repository at https://repositorio.ipea.gov.br/bitstream/11058/13124/1/Os_alunos_e_as_escolas_no_Brasil_de_hoje.pdf Fletcher and Moura Castro, at the time this work was published, were members of several national and international bodies, for example UNDP – the United Nations Development Programme – and the ILO – International Labour Organization –, both of which still exist within the United Nations system. In addition, they were part of IPLAN – the Institute of Planning –, no longer in existence, but which was then part of IPEA – the Institute for Applied Economic Research –, today part of the Ministry of Planning and Management. Costa Ribeiro, SERGIO. *A pedagogia da repetência*. Estudos Avançados, volume 5, no. 12, USP. São Paulo, May/August 1991.
19. Harbison, Ralph W.; Hanushek, Eric A. *Educational performance of the poor: lessons from rural northeast Brazil* (English). New York: Oxford University Press. <http://documents.worldbank.org/curated/en/310981468770471105>
20. Cf. SAEB. Histórico. See at <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb/historico> Cf. Novas Perspectivas. Brasília, April 2002. See at https://download.inep.gov.br/download/saeb/2001/Miolo_Novas_Perspectivas2001.pdf
21. There are instructivist pedagogical approaches to the problem of teaching and learning. There is also a psychometric approach based on KST – Knowledge Space Theory –, rather than on IRT. Cf. Nature Research Intelligence Topics. Knowledge Space Theory and Skill Assessment. See at <https://www.nature.com/research-intelligence/nri-topic-summaries/knowledge-space-theory-and-skill-assessment> With regard to instructivist approaches, or at least to non-constructivist approaches, see: Mayer, R. E. Should There Be a Three-Strikes Rule Against Pure Discovery Learning? *American Psychologist*, Volume 59, Issue 1, 14–19, 2004. Clermont GAUTHIER, M'hamed MELLOUKI, Denis SIMARD, Steve BISSONNETTE and Mario RICHARD. Quelles sont les pédagogies efficaces? Un état de la recherche. *Les Cahiers du débat*. Fondation pour l'innovation Politique, January 2005. Cf. Paul A. Kirschner, John Sweller & Richard E. Clark. Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching, *Educational Psychologist*, Volume 41, Issue 2, 75–86, 2006. Cf. Stockard, J., Wood, T. W., Coughlin, C., & Rasplika Khoury, C. The Effectiveness of Direct Instruction Curricula: A Meta-Analysis of a Half Century of Research. *Review of Educational Research*, Volume 88, Issue 4, 479–507, 2018. Juan C. Castro-Alonso, Alejandro A. Hidalgo and John Sweller. Biological evolution and human cognition are analogous information processing systems. *Frontiers in Psychology*, Volume 14, 2023. Regarding the mathematical modeling used to interpret psychometric test data by means of KST – Knowledge Space Theory –, see: Falmagne, J-C; Doignon, J-P, Koppen, Mathieu; Villano, Michael; Leila Johannesen. Introduction to Knowledge Spaces: How to Build, Test, and Search Them. *Psychological Review*, Vol. 97. No. 2, 201–224, 1990. The authors hold the following academic affiliations; the authors' names are given in full. Falmagne, J-C et al. The Assessment of Knowledge in Theory and in Practice. In: *Formal Concept Analysis*. 4th International Conference, ICFA 2006, Dresden, Germany, February. Conference proceedings. Masura, Amanda. Knowledge Structure Evaluation of ENEM: do different groups learn differently? Master's thesis submitted to the Stricto Sensu Graduate Program in Psychology at Universidade São Francisco.
22. See the entries for Item Response Theory and Latent Trait Theory in the American Psychological Association's Dictionary of Psychology: <https://dictionary.apa.org/item-response-theory> and <https://dictionary.apa.org/latent-trait-theory>
23. At this point, in a technical register generally belonging to measurement theory, and in particular to the psychometric and mathematical fields involved in the statistics describing the distribution of numerical proficiency values on tests, it becomes necessary to discuss the concept of the Z score, or standard score. This means: instead of looking at the numerical difference between test scores – the raw score – the psychometrician, knowing that what is very likely being measured is a normally distributed ability within the population of interest – few cases with very high or very low scores and most with neither very high nor very low values – finds it more informative to look at the distance between a given score and the calculated mean of all scores. The Z score indicates how far a score lies, above or below, from that calculated mean value, which in turn indicates the numerical proficiency value around which most values are concentrated. The unit in which the Z score is expressed is called the standard deviation. There is a mathematical formula for calculating it. This way of looking at scores is where the harmonization between the pedagogical approach and the psychometric approach adopted in SAEB is realized – socio-interactionist cognitivism and IRT. Whereas for the former, teaching/learning experiences never result in nullity nor in the full development of cognitive competencies/skills, for the psychometric approach, although the probability of a correct answer to an item increases or decreases according to whether the associated latent trait is greater or smaller, that probability is never 100% nor 0%, however great or small the latent trait may be. What matters is to observe average performance closely and to compare individual performances against it. The understanding is that efforts in education policy should focus on raising that average performance.
24. Cf. Mark D. Reckase. *A Tale of Two Models: Sources of Confusion in Achievement Testing*. ETS Research Report Series. Volume 2017, Issue 1. Pages 1–11.
25. This means that any member of that school population, chosen at random, even if they had not taken the test, would if they did have a high probability of performing in line with the average performance of the population that did take it.
26. To interpret a test proficiency pedagogically is to describe what the students whose proficiency values fall within each level or numerical range into which the scales are divided know and are able to do. This interpretation is made separately for each cognitive domain in which the tests are administered.
27. MEC. INEP. *Relatório Nacional Saeb 2003*. Brasília, 2006. Pages 45 to 51. Available for reading at https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/relatorio_nacional_do_saeb_2003.pdf
28. MEC/INEP. *Qualidade da Educação: uma nova leitura do desempenho dos estudantes da 4ª série do ensino fundamental*. Brasília, April 2003, page 7. https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/qualidade_da_educacao_uma_nova_leitura_do_desempenho_dos_estudantes_da_4_serie_do_ensino_fundamental.pdf Available for reading. Accessed on 23/06/2025.
29. *Ibid*, pages 7 to 11. MEC/INEP. *Qualidade da Educação: uma nova leitura do desempenho dos estudantes da 8ª série do Ensino Fundamental*. Brasília, December 2003, pages 9 and 10. Available in the INEP publications repository. Accessed on 23/06/2025. MEC/INEP. *Qualidade da Educação: uma nova leitura do desempenho dos estudantes da 3ª série do Ensino Médio*. Brasília, January 2005, pages 10 and 11. Available for reading at https://download.inep.gov.br/download/saeb/2004/qualidade_educacao.pdf Accessed on 23/06/2025.

30. MEC/INEP. SAEB - 2005 Primeiros Resultados: Médias de desempenho do SAEB/2005 em perspectiva comparada. Brasília, February 2007. Available for reading at https://download.inep.gov.br/educacao_basica/prova_brasil_saeb/resultados/SAEB1995_2005.pdf In 2017, there was an attempt to report results in a way similar to that used in 2003 and 2004. In that attempt, however, performance classification was reduced to three levels: insufficient, basic, and adequate. But it never even made it into the final SAEB 2017 report. It appeared only in the Press Kit, a summary of results for release to the press ahead of publication of the report with the SAEB data. See MEC. INEP. Assessoria de Comunicação Social. Press Kit SAEB 2017. Brasília. Page 15. Available for reading at https://download.inep.gov.br/educacao_basica/saeb/2018/documentos/presskit_saeb2017.pdf
31. From 2017 onward, INEP began releasing, only as an Excel file on INEP's website, under the title Planilha de Resultados Brasil, information on the percentage of proficiencies by SAEB scale level. But this is of limited pedagogical use without also releasing the minimum proficiencies expected for each stage of schooling in each cognitive domain tested. See at https://download.inep.gov.br/educacao_basica/saeb/2019/resultados/Resultados_Saeb_2017_Brasil_Estados_Municipios.zip
32. The socioeconomic context behind the results obtained is, without doubt, one criterion for judging them. But should it be the main, predominant one? INEP's reports carry out result analyses that lead to comparing a given unit's average proficiency with the national average. That would be one criterion for judgment. But they also foster the habit of favoring, in that comparison, matching results between educational units with similar profiles in terms of student socioeconomic indicators, teacher qualification levels, and administrative management effort.
33. See, in note 15, the trend in proficiencies across the various stages of schooling.
34. IBGE. Estudos & Pesquisas. Informação Demográfica e Socioeconômica. No. 36. Síntese de Indicadores Sociais. Uma Análise das Condições de Vida da População Brasileira 2016. Chart 4.2.
35. See the Ideb intermediate targets at <https://app.powerbi.com/view?r=eyJrIjojMGVjMzlwZWQzM2IzZS00NmE0LTkwNjU0ZjI1YjMyNTVhZGY0IiwidCI6IjI2Zjc0dK3LWM4YWVhNGVhZS05NzhmLWVhNGMwNzc0MzRiZjIj>
36. Cf. Ministério da Educação. Brasil. Governo Federal. PDE – O Plano de Desenvolvimento da Educação. Razões, princípios e programas. Brasília, 2007. Page 21. See at <https://portal.mec.gov.br/arquivos/livro/livro.pdf> The financial costs of persistently high grade-repetition rates are large. The association between repetition and a dysfunctional school trajectory is a constant in the literature on the subject. For ethical reasons, controlled experiments – which would have greater power to clarify whether this association is causal, and in which direction – are not feasible in this case. The balance of benefits and risks of grade retention therefore remains inconclusive. The use of mathematical tools in variable analysis attempts to reduce the risk of selection bias in purely observational studies that examine the associations between repetition, grade delay, dropout, and so on. But the findings of these studies increasingly discredit the all-or-nothing approach. See: Jane Cooley Fruehwirth, Salvador Navarro, and Yuya Takahashi. How the Timing of Grade Retention Affects Outcomes: Identification and Estimation of Time-Varying Treatment Effects. Journal of Labor Economics Volume 34, Number 4 October 2016. Emilio Borghesan, Hugo Reis, and Petra E. Todd. Learning Through Repetition? A Dynamic Evaluation of Grade Retention in Portugal. PIER. Penn Institute for Economic Research. Working Paper 22-030. 2022. Emilio Borghesan, Hugo Reis, and Petra E. Todd. If at First You Don't Succeed: A Dynamic Evaluation of Grade Retention. Paper presented in seminars at U Penn and Brandeis University, SOLE 2022 Conference, and as a keynote lecture at the Brazilian Econometric Society meetings and the Barcelona School of Economics Summer Forum. February 19, 2025.
37. IPEA. Políticas Sociais. Acompanhamento e Análise. No. 16. 2008. Page 108. See at https://portalantigo.ipea.gov.br/portal/images/stories/PDFs/politicas_sociais/bps_16_completo.pdf
38. Brazil: Encouraging Lessons from a Large Federal System. In OECD, Lessons from PISA for the United States. Strong Performers and Successful Reformers in Education, 2011. OECD Publishing. <http://dx.doi.org/10.1787/9789264096660-en> OECD. Education and Skills YouTube Channel. 2012. Brazil - Strong Performers and Successful Reformers in Education. See at <https://youtu.be/Sm6zAMemfYs> Barbara Bruns, David Evans, Javier Luque. Achieving World-Class Education in Brazil. The Next Agenda. World Bank. Directions in development. Human development, 2012. See at <https://openknowledge.worldbank.org/server/api/core/bitstreams/4f1f0074-1e4d-559c-a347-1a9294db2918/content>
39. Methodological note on aligning PISA performance with the SAEB scale. Available for reading at https://download.inep.gov.br/educacao_basica/portal_ideb/o_que_sao_as metas/Nota_Tecnica_n3_compatibilizacao_PISA_SAEB.pdf
40. The pedagogical interpretation of the 8- and 10-interval scales was established in the SAEB 2001 reports and kept in the 2003 edition. Cf. INEP. Relatório SAEB 2001. Língua Portuguesa. Page 14. Brasília, 2002. Available for reading at https://download.inep.gov.br/download/saeb/2001/relatorioSAEB_portugues.pdf INEP. Relatório SAEB 2001. Matemática. Page 14. Brasília, 2002. Available for reading at https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_examenes_da_educacao_basica/relatorio_saeb_2001_matematica.pdf
41. Between 2001 and 2023, SAEB used proficiency scales referenced to the Matrizes de Referência (Reference Matrices) drawn up in 2001. Throughout this period, its test items were developed based on those Reference Matrices. Starting in 2025, the proficiency scales will be based on Reference Matrices updated in line with the BNCC. This update has so far only been carried out for the elementary-education SAEB tests, 2nd, 5th, and 9th grades. Cf. the SAEB 2001 Reports for Portuguese Language and Mathematics referenced in the previous note. BRASIL. MEC. INEP. DAEB. Escalas de Proficiência do SAEB. Brasília, August 2020 (edition superseded, in the INEP repository, by the 2024 version cited below). BRASIL. MEC. INEP. DAEB. Escalas de Proficiência do SAEB. 2nd, 5th and 9th grade of elementary education. Brasília, 2024. See at https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_examenes_da_educacao_basica/escalas_de_proficiencia_do_sae_b.pdf
42. OECD. Learning for Tomorrow's World First Results from PISA 2003. Page 45 and page 273. See at https://www.oecd.org/content/dam/oecd/en/publications/reports/2004/12/learning-for-tomorrow-s-world_g1gh487c/9789264006416-en.pdf OECD. PISA 2022 Results. The State of Learning and Equity in Education Publication. Page 89 and page 98. Available for reading at https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/12/pisa-2022-results-volume-i_76772a36/53f23881-en.pdf
43. OECD. Learning for Tomorrow's World First Results from PISA 2003. Table 2.5a, page 354 and Table 6.1, page 443. See at https://www.oecd.org/content/dam/oecd/en/publications/reports/2004/12/learning-for-tomorrow-s-world_g1gh487c/9789264006416-en.pdf OECD. PISA 2022 Results. The State of Learning and Equity in Education Publication. Table I.B1.3.1, page 317 and Table I.B1.3.3, page 321. Available for reading at https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/12/pisa-2022-results-volume-i_76772a36/53f23881-en.pdf
44. OECD. Learning for Tomorrow's World First Results from PISA 2003. Table 2.5a, page 354 and Table 6.1, page 443. See at https://www.oecd.org/content/dam/oecd/en/publications/reports/2004/12/learning-for-tomorrow-s-world_g1gh487c/9789264006416-en.pdf
45. OECD. Learning for Tomorrow's World First Results from PISA 2003, page 92, Figure 2.16b (Mathematics) and page 281, Figure 2.3 (Reading). Available for reading at https://www.oecd.org/content/dam/oecd/en/publications/reports/2004/12/learning-for-tomorrow-s-world_g1gh487c/9789264006416-en.pdf Note, on page 33 of the report linked above, the following caveat: "In this publication, the OECD total is generally used when references are made to the stock of human capital in the OECD area. Where the focus is on comparing performance across education systems, the OECD average is used."

46. INEP. Nota metodológica sobre a compatibilização de desempenhos do PISA com a escala do SAEB. Table 1, page 2. See at https://download.inep.gov.br/educacao_basica/portal_ideb/o_que_sao_as metas/Nota_Tecnica_n3_compatibilizacao_PISA_SAEb.pdf OECD. Learning for Tomorrow's World First Results from PISA 2003, Table 2.5c, page 356 and Table 6.2, page 444. Available for reading at https://www.oecd.org/content/dam/oecd/en/publications/reports/2004/12/learning-for-tomorrow-s-world_g1gh487c/9789264006416-en.pdf
47. OECD. Learning for Tomorrow's World First Results from PISA 2003, page 58, Box 2.1, page 92, Table 2.16b, and page 281, Table 6.3. Available for reading at https://www.oecd.org/content/dam/oecd/en/publications/reports/2004/12/learning-for-tomorrow-s-world_g1gh487c/9789264006416-en.pdf
48. In general, the appropriate grade for someone aged 15 in 2003 was the 8th grade of elementary education, or the 1st year of secondary school. For someone aged 16, the 1st or 2nd year of secondary school. Today these grades would be the 9th year of elementary education/1st year of secondary school, or the 1st/2nd years of secondary school, for ages 15 and 16 respectively. In Brazil, in 2003, the actual age-grade distortion rate in the population aged 15 and 16 was likely higher than the age-grade distortion rate in that sample. It is known that proficiency tends to be higher the lower the age-grade distortion. Precise information on the difference between the age-grade distortion rate in the 15-to-16-year-old population in 2003 and the age-grade distortion rate of the Brazilian PISA 2003 sample would require analysis of microdata from both the 2003 School Census and PISA 2003. This is made even harder because Educacenso, which brought greater precision to School Census data, was only introduced in 2007. In any case, there is a strong likelihood that the Brazilian population tested in PISA 2003 had a better age-grade profile than the Brazilian population aged 15 and 16 at the time. Data on the Brazilian sample in PISA 2003 are on pages 320-322 of that edition's report. See Table A3.1, columns 1 and 2.
49. OECD. PISA. Measuring Student Knowledge and Skills. 1999. Table 1. Summary of PISA dimensions. Page 11. See at <https://doi.org/10.1787/9789264173125-en> MEC/INEP. Saeb 2001. Novas perspectivas. Brasília, April 2002. Pages 9 to 29. See at https://download.inep.gov.br/download/saeb/2001/Miolo_Novas_Perspectivas2001.pdf
50. INEP. NOTA TÉCNICA. Metodologia utilizada para o estabelecimento das metas intermediárias para a trajetória do Ideb no Brasil, Estados, Municípios e Escolas. Page 5. Available for reading at https://download.inep.gov.br/educacao_basica/portal_ideb/o_que_sao_as metas/Nota_Tecnica_n2_metas_intermediarias_IDEB.pdf Accessed on June 14, 2025.
51. INEP's own data report the following proportions of students who are not literate by the end of the 2nd year of elementary education, all with 4 years of schooling, by school network: Public and Private networks: 56%. Public network, State and Municipal: 64%. Private network: 21%. See at https://download.inep.gov.br/saeb/resultados/saeb_2021_brasil_estados_municipios_c_tx_alfabetizado.xlsx Regarding the trend in PIRLS and PISA proficiencies, see note 15.
52. Although it has fallen 10% since 2016, the percentage of Brazilians who have not completed basic education stood at 44% in 2024. It appears, then, that despite sustained high enrollment rates over the last decade, between 2016 and 2024, Brazilian education seriously fails to bring the school-age population to the completion of basic education. This is what the Painel Educação of IBGE's Pnad Continua shows. Enrollment rates evolved, between 2016 and 2024, as follows: from 90.0% to 92.9% for the 4-and-5-year-old group; 99.2% to 99.4% for the 6-to-14 group; 86.9% to 91.9% for the 15-to-17 group. Cf. IBGE. PNAD Continua. Educação. 2024. See at <https://painel.ibge.gov.br/pnadc/> By the INAF – National Literacy Indicator – criterion, the following are considered functionally illiterate in Brazil: 29% of the youth and adult population, 17% of those with a secondary-school diploma, and 12% of those with higher education. Cf. INAF 2024. Cf. INAF 2024. Indicador de Alfabetismo Funcional. See at <https://alfabetismofuncional.org.br/dados-inaf-2024.pdf>
53. INSTITUTO UNIBANCO. Saeb e Ideb. Recomendações para Interpretar os Resultados. Nota Técnica. Undated. Available for reading at <https://observatoriodeeducacao.institutounibanco.org.br/api/assets/observatorio/b094a67c-2ed8-42dc-a7ea-c2094ff0b29d/instituto-unibanco-notatecnica-ideb2023-final.pdf?version=0> IEDE. Interdisciplinaridade e Evidências no Debate Educacional. Propostas e Reflexões para o Novo IDEB. August 2022. Available for reading at https://portaliede.org.br/wp-content/uploads/2022/08/Propostas_Reflexoes_Novo_Ideb_Agosto2022-2.pdf
54. On CAQ and CAQi, see note 12. On INEP's technical notes, they are available for reading at <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/fundeb/notas-tecnicas>
55. In the United States there is a National Academies system made up of 3 academies: the National Academy of Sciences itself, plus two others, the National Academy of Engineering and the National Academy of Medicine.
56. National Research Council. Uncommon Measures: Equivalence and Linkage Among Educational Tests, 1999. Washington, DC: The National Academies Press. American Educational Research Association. American Psychological Association. National Council on Measurement in Education. Standards for Educational and Psychological Testing, 2014. Michael J. Kolen and Robert L. Brennan. Test Equating, Scaling, and Linking Methods and Practices. Third Edition, Springer, 2014.

INSTITUTO
ARRECIFE



[ADDRESS TO BE CONFIRMED]

