

ASSESSING TO CARE: *formative assessment, authorship, and the monitoring of learning¹*

Pedro Demo²

Silvane Friebe³

Cristiano de Souza Calisto⁴

ABSTRACT

This essay discusses educational assessment as a diagnostic, formative, and caring practice, in contrast to its bureaucratic and instructionist use. It aims to critically interpret how assessment has been conceived and implemented in schools, particularly when it becomes disconnected from learning. It is characterized as a theoretical-interpretive essay with a qualitative approach, guided by Demo's (2001) reconstructive perspective and supported by bibliographic and documentary analysis of public data from the 2024 ENEM and the 2023 Ideb. Diagnosis is understood in a pedagogical sense, as a process-based, contextual, and revisable mapping of learning, intended to guide decisions and interventions without pathologizing difficulties, differences, or errors. The analysis highlights the dissociation between assessment and learning and affirms the centrality of teacher follow-up, feedback, rewriting, authorship, and process records. It concludes that assessing in order to care means transforming evidence into the reorganization of pedagogical work and

¹ English version by Aldeisa Santos de Carvalho *E-mail:* aldeisadoutorado@gmail.com.

² PhD em Sociologia pela Universität des Saarlandes, Saarbrücken (Alemanha, 1971). Professor Doutor Titular do Departamento de Sociologia da UnB. Docente voluntário no Programa de Pós-Graduação em Direitos Humanos e Cidadania (PPGDH) da Universidade de Brasília. Orcid: <https://orcid.org/0000-0002-5514-2781>. *E-mail:* pedrodemo@gmail.com.

³ Doutoranda e Mestre em Direitos Humanos e Cidadania pelo Programa de Pós-graduação em Direitos humanos e Cidadania da Universidade de Brasília (UnB). Orcid: <https://orcid.org/0009-0000-4201-5217>. *E-mail:* silvanefriebe@gmail.com.

⁴ Doutorando em Direitos Humanos e Cidadania pelo Programa de Pós-graduação em Direitos humanos e Cidadania da Universidade de Brasília (UnB). Orcid: <https://orcid.org/0009-0006-5655-1305>. *E-mail:* cristianocalisto@gmail.com.

into new learning opportunities, including in the face of the challenges posed by the digital world and Artificial Intelligence.

KEYWORDS: Formative Assessment. Pedagogical Diagnosis. Reconstructive Perspective. Authorial Learning. Pedagogical Care.

AVALIAR PARA CUIDAR: avaliação formativa, autoria e acompanhamento das aprendizagens

RESUMO

Este ensaio discute a avaliação educacional como prática diagnóstica, formativa e cuidadora, em contraste com seu uso burocrático e instrucionista. Objetiva interpretar criticamente como a avaliação tem sido concebida e operacionalizada na escola, sobretudo quando se distancia da aprendizagem. Caracteriza-se como ensaio teórico-interpretativo, de abordagem qualitativa, orientado pela perspectiva reconstrutiva de Demo (2001) e articulado à análise bibliográfica e documental de dados públicos do ENEM 2024 e do Ideb 2023. O diagnóstico é compreendido em sentido pedagógico, como mapeamento processual, contextual e revisável das aprendizagens, destinado a orientar decisões e intervenções, sem patologizar dificuldades, diferenças ou erros. A análise evidencia a dissociação entre avaliação e aprendizagem e sustenta a centralidade do acompanhamento docente, das devolutivas, da reescrita, da autoria e dos registros processuais. Conclui-se que avaliar para cuidar significa transformar evidências em reorganização do trabalho pedagógico e em novas possibilidades de aprendizagem, inclusive diante dos desafios do mundo digital e da Inteligência Artificial.

PALAVRAS-CHAVE: Avaliação Formativa. Diagnóstico Pedagógico. Perspectiva Reconstrutiva. Aprendizagem Autoral. Cuidado Pedagógico.

EVALUAR PARA CUIDAR: evaluación formativa, autoría y acompañamiento de los aprendizajes

RESUMEN

Este ensayo analiza la evaluación educativa como una práctica diagnóstica, formativa y cuidadora, en contraste con su uso burocrático e instruccionalista. Su objetivo es interpretar críticamente cómo la evaluación ha sido concebida y aplicada en la escuela,

especialmente cuando se distancia del aprendizaje. Se caracteriza como un ensayo teórico-interpretativo, de enfoque cualitativo, orientado por la perspectiva reconstructiva de Demo (2001) y articulado con el análisis bibliográfico y documental de datos públicos del ENEM 2024 y del Ideb 2023. El diagnóstico se comprende en un sentido pedagógico, como un mapeo procesual, contextual y revisable de los aprendizajes, destinado a orientar decisiones e intervenciones, sin patologizar las dificultades, las diferencias ni los errores. El análisis evidencia la disociación entre evaluación y aprendizaje y sostiene la centralidad del acompañamiento docente, la retroalimentación, la reescritura, la autoría y los registros procesuales. Se concluye que evaluar para cuidar significa transformar las evidencias en reorganización del trabajo pedagógico y en nuevas posibilidades de aprendizaje, incluso ante los desafíos del mundo digital y de la Inteligencia Artificial.

PALABRAS CLAVE: Evaluación Formativa. Diagnóstico Pedagógico. Perspectiva Reconstructiva. Aprendizaje Autoral. Cuidado Pedagógico.

* * *

Introduction

Persistent low performance in basic and higher education has not led, as might be expected, to the strengthening of an assessment culture committed to learning. Instead, school failure has become normalized, while an outdated instructionist model persists, centered on the lesson as an end in itself and largely unresponsive to what students actually learn (Frigotto, 1988; 2018; Libâneo, 2012). Recent indicators make this situation especially stark: in the 2024 National High School Examination (Exame Nacional do Ensino Médio, ENEM), only 12 participants achieved the maximum essay score among approximately 3.2 million registered candidates, a statistically negligible proportion (Casemiro and Santos, 2025). Similarly, data from the 2023 Basic Education Development Index (Índice de Desenvolvimento da Educação Básica, Ideb) show that, in public upper-secondary education, the

proportion of students meeting the Adequate Learning benchmark in mathematics ranged from 2% to 9% across the states, with a national average of 5%. In private schools, the corresponding figures ranged from 15% to 49%, with an average of 31% - still an insufficient level, particularly given the socioeconomic profile of the population served and the fact that this is fee-paying provision (QEdu, 2023).

Beyond the national picture, progression across educational stages reveals a sharp decline in mathematics, particularly in the transition to the later years of elementary school and then to upper-secondary education. This decline calls into question the rationale underlying existing structures for pedagogical monitoring. From an educational-policy perspective, the evidence suggests that assessment, although widely institutionalized, has rarely been mobilized as an instrument of diagnosis and care. In many cases, it functions more as a bureaucratic ritual than as formative mediation aimed at reducing inequalities and reconstructing learning trajectories. In the Federal District, for example, the longitudinal decline in mathematics and the deterioration between educational stages underscore both the severity of the problem and the inadequacy of an assessment culture centered on classroom instruction.

Part of the impasse appears to lie in the refusal of diagnosis because of its political and symbolic severity: when diagnosis is avoided, an appearance of normality is preserved, even in the face of results that should trigger structural change. Darcy Ribeiro's well-known assertion that public schooling is not in crisis but is itself a project helps explain why, in contexts marked by institutional accommodation, "assessment" may become a strategy of concealment (Roitman, 2022). In this essay, diagnosis is used in a pedagogical rather than clinical sense. From Luckesi's (2011) perspective, assessment entails interpreting relevant information about learning in order to support decisions and redirect pedagogical action. Diagnosis therefore does not confine the learner to a definitive classification; instead, it offers a provisional reading of the learning situation designed to guide intervention. From the same perspective, error may be understood as an indicator of the learning

trajectory and a resource for further learning, rather than as grounds for punishment or a marker of incapacity (Luckesi, 1990).

This understanding requires recognition of the limits involved in transferring medical categories to the educational field. Medicalization occurs when difficulties produced within the schooling process are converted into symptoms or individual problems of a predominantly biological nature, while the pedagogical, institutional, historical, and social conditions that shape learning are disregarded (Meira, 2012; Scarin; Souza, 2020). Accordingly, in this article, pedagogical diagnosis refers to a process-based and revisable mapping of learning, the interpretation of learning trajectories, and the identification of difficulties and potentialities in order to guide monitoring and intervention, without pathologizing error, difference, or non-learning.

Against this background, the study critically examines how assessment has been conceived and operationalized in schools, particularly when it becomes detached from its formative and care-oriented function. The methodological approach and the criteria used to select and analyze the materials are presented in the following section.

Methodological Approach

This study is a theoretical-interpretive essay with a qualitative approach, guided by Demo's (2001) reconstructive perspective. It combines bibliographic and documentary analysis, without claiming to constitute a systematic review or to produce statistical generalizations.

The bibliographic corpus comprised works related to the following analytical axes: educational assessment; pedagogical diagnosis and critiques of medicalization; authorial learning; critiques of instructionism; pedagogical monitoring; formative and mediating assessment; assessment culture; educational inequalities; relations among assessment, classification, and power; and educational challenges associated with the digital world and artificial intelligence. References were selected

purposely on the basis of their relevance to the research problem, their conceptual contribution to the argument, and their capacity to establish contrasts between classificatory assessment practices and formative practices oriented toward monitoring learning.

The documentary analysis drew primarily on publicly available educational data and indicators from the 2024 ENEM and the 2023 Ideb, supplemented by information from other sources cited in the manuscript, such as the 2025 Functional Literacy Indicator (Indicador de Alfabetismo Funcional, Inaf). The indicators were selected to illuminate: (a) the performance of examination participants and students; (b) differences between public and private school systems; (c) the progression of results across stages of basic education; and (d) specific contexts, such as the results observed in the Federal District. These data were used as contextual evidence rather than subjected to an original statistical analysis.

The analytical procedure articulated three complementary and not necessarily linear movements: identifying evidence of low performance and educational inequality; confronting this evidence with the literature on assessment, learning, authorship, and instructionism; and interpreting it through the categories of pedagogical diagnosis, monitoring, feedback, rewriting, authorship, care, and the reorganization of pedagogical work.

The documents and indicators were used not merely to describe outcomes, but to support a critical analysis of the disjunction between assessment and learning and to substantiate the case for process-based, formative, dialogical, and care-oriented assessment practices.

Challenges in Assessing Learning

External assessments provide useful reference points, although they remain removed from everyday school life. Determining whether students are learning is a highly complex challenge. Systems such as the Programme for International Student Assessment (PISA),

administered by the Organisation for Economic Co-operation and Development (OECD), and Brazil's Ideb generate relevant large-scale information, but they do not replace assessment conducted in the daily life of schools. Everyday assessment is central to organizing pedagogical interventions and relies on continuous diagnostic monitoring (Friebel, 2025). The use of established technical models, such as Item Response Theory (IRT), facilitates comparison, although applying them in a country of continental dimensions and profound inequalities requires caution. Moreover, the very concept of “adequate learning” remains contested, because attempts to measure learning tend to privilege quantifiable dimensions, especially content mastery. The problem also lies in the isolated use of this information. Results from external assessments must be brought into dialogue with assessment conducted in everyday school settings so that they can inform the replanning of pedagogical work and the monitoring of learning (Villas Boas; Dias, 2015).

Authorship as an assessment criterion, however, requires the quality of students' authorial writing to be appraised by the teacher through close and sustained engagement during its development. A glimpse of this appears in the ENEM, which requires an essay - a fundamental demand, even though the prescribed format is highly formulaic, partly to make large-scale marking operationally manageable. The result is telling: writing remains an exceptional form of elaboration in schools, even though any robust theory of deep learning requires sustained authorial practice (Demo, 2015; 2018). “Learning appears to occur only when the learner pays attention, thinks, anticipates, and sets hypotheses in motion while risking error” (Demo; Friebel, 2025, p. 107). With regard to the limits of automated assessment and the risk of simulation, there are theoretical and methodological dimensions that assessments such as PISA and Ideb cannot capture. They therefore cannot replace assessment carried out by a teacher in interaction with the student. If learning entails authorial elaboration and the production of meaning, its assessment requires more substantial and

recurrent forms of production through which the student's creative capacity can emerge (Dehaene, 2022).

Closed-response questions reveal little about authorship, both because answers may be correct by chance and because such questions do not make the open process of elaboration visible. In assessments administered to millions of students, automated marking may be operationally necessary; nevertheless, it narrows the range of what can be observed. Authorship requires original, non-standardized, and non-plagiarized work, assessed by a teacher who follows the process and sustains a deep, reciprocal, intersubjective relationship with the student.

For process-based assessment involving digital records, rewriting, and progression, one possibility is to assemble students' texts and the feedback they receive in a digital portfolio, preserving them for future comparison. This procedure presupposes that students produce texts systematically - monthly, weekly, or daily, depending on the context. Unlike the ENEM, in which many students confront an "essay" only on the day of the examination because writing is treated as an exceptional procedure, the assessment of authorial learning requires regular, ideally daily, writing as a structural mode of learning: learning through elaboration. Students engage systematically in learning activities under attentive and continuous teacher guidance. They write not only to be assessed, but to learn as authors. Organizing these productions in a digital portfolio is consistent with Villas Boas's (2004) conception of the portfolio as a formative assessment procedure that enables the monitoring of learning development and fosters reflection, participation, and autonomy. This practice also resonates with Hoffmann's (2012) conception of mediating assessment, in which assessment ceases to be a terminal act of classification and becomes continuous monitoring attentive to each learner's possibilities for progress.

By comparing, for example, a text written in April with one written in March, it should be possible to identify progress through qualitative indicators such as denser reading and counter-reading, more intensive

study, more penetrating inquiry, more tangible originality, focused argumentation, and clearer, more vigorous communication. The digital portfolio provides a continuing record of the student's trajectory toward greater authorial competence and may also be accessed electronically by those responsible for the learner. What matters is not the grade itself, but the student's authorial process and product. Thus, when a text is not yet adequate, the student can rewrite it, under teacher guidance, until it reaches an appropriate standard. By examining the records gathered in the portfolio, learners can assess their own progress and assume greater responsibility for their development, thereby strengthening intellectual autonomy (Villas Boas, 2006). The central question is not merely whether lessons were delivered, but whether students are progressively developing authorship; whether they are learning to read and counter-read an author in order to become authors themselves; whether they can engage with science, research, and textual elaboration; whether they can use research methods and techniques competently; whether their command of textual grammar is improving; and whether they can construct theory of their own, which is ultimately what matters most.

Understanding assessment as a qualitative and dialectical process of continuing formation does not entail disregarding quantitative data (Demo; Calisto, 2025). Assessment is an open and continuous process, as is authorship in formation. Because there is no definitive text and no finished author, students' productions are essays: they remain open to reformulation and critique, a fundamental condition of their scientific character. The purpose of writing is to foster authorial learning.

In the digital world and in the context of artificial intelligence (AI), the distinction between syntax and semantics becomes even more consequential. AI operates primarily at the syntactic level: it combines formal patterns, organizes sentences, and reproduces structures with great efficiency, but it does not, by itself, address the semantic dimension of learning, which involves authorship, intention, lived experience, and the contestation of meaning. School assessment therefore cannot be reduced to content checking or the evaluation of

“ready-made” products; it must focus on the student's authorship as revealed through the process of elaboration. This reinforces the centrality of assessment practices involving feedback, rewriting, and monitoring - for example, through a digital portfolio - in which texts are revisited, improved, and assumed as authorial constructions. For such monitoring to produce learning, feedback cannot be limited to identifying correct and incorrect responses. Villas Boas (2006) regards feedback as an essential component of formative assessment, particularly when it contributes to self-monitoring and intellectual autonomy. Similarly, Hattie and Timperley (2007) show that the effects of feedback depend on both its content and the manner in which it is provided: it should clarify learning goals, the progress achieved, and the next steps required. Understood in this way, feedback guides revision and reconstruction rather than merely justifying a grade.

With regard to the relationships among authorship, literacy, and democracy, preventing students from producing texts of their own perpetuates a form of illiteracy marked by copying rather than elaboration. Human production always draws on prior repertoires, yet it can acquire singularity, style, and originality. AI may be misused when it serves only automatic reproduction; even so, it can provide material for original elaboration, provided that it is used critically rather than submissively (Varoufakis, 2023).

It follows that assessment should focus on students' textual competence, now also in digital environments, which can vastly expand their possibilities or reduce them to systematic plagiarism. Writing involves more than assembling graphic symbols, which are merely instrumental. It entails formulating working hypotheses (Piaget, 1996), testing them in laboratories and experiments, theorizing with confidence, constructing one's own theory from existing theories, learning to argue and substantiate claims, and, above all, learning to disagree productively and respectfully. It also requires knowing how to read and, especially, how to counter-read, because we read authors in order to become authors ourselves, not their spokespersons. As a syntactic artifact, a text must express the student's semantic capacity

by articulating ideas, meanings, original insights, and formulations of their own. At the semantic level, ideas are singular because they arise from lived experience; they are therefore inherently divergent and generate the diversity of meanings that is crucial to human development. In every text, grammar is shared, but lived experience is not. We may focus on what is repeated and grammatical, or we may focus on what is divergent, singular, and authorial. Schools, like universities, frequently privilege content mastery at the expense of assessing authorship and processes of elaboration.

Thinking, disagreeing, arguing, and theorizing with intellectual ownership often appear uncomfortable, challenging, and disruptive - activities that institutions may prefer to suppress. It is easier to play with a parrot that stammers or repeats words but neither challenges its interlocutor nor contests meanings nor aspires to citizenship. The fact that so few students can write proficiently, as the ENEM so starkly reveals, suggests a student population composed largely of trained parrots. In terms of functional literacy, only 10% are classified as “proficient,” according to the Functional Literacy Indicator (Inaf, 2025), a proportion similar to that found for digital literacy. This deprivation directly compromises the democratic quality of the population. “Knowing how to think” has yet to reach schools and universities. What value can “functional literacy” possibly have? Its value lies in moving beyond the mechanical command of words and grasping how they become saturated with contested, diverse, and creative meanings. We then discover that, whereas approximately 7% of the population is conventionally “illiterate,” 29% is “functionally illiterate.” This is what remains of such an inept school system (Demo; Calisto, 2025a).

Assessment Is Always Contestable

The argument begins from the premise that no assessment can be entirely fair, because assessment is always conducted by human beings, who are themselves imperfectly just. Even the judge, one of

society's most institutionally recognized evaluators, cannot be fully "just," because judgment remains shaped by standpoint, however informed or well trained the judge may be in judicial procedure. Moreover, judging is always contentious in social life and may even be bought and sold - as in a market for judicial rulings. Every minimally reflective teacher knows that assessments are approximate, risky, and readily contestable. No human judgment is entirely reliable because no human mind is entirely reliable. Judgments are "self-referential": they do not apprehend reality directly from reality itself, whose nature we do not know (Hoffman, 2019), but through our own mental projections. We do not see things as they are, but as we are. This does not condemn us to ideological closure, because social life exposes our viewpoints to others and allows us to recalibrate them intersubjectively. Yet we inevitably see through our own eyes, not through those of others. Judges must learn to be "impartial," but if they are genuinely reflective, they recognize their own partiality. Complete impartiality would require shedding subjectivity, which is impossible. We cannot step outside our own skin to observe ourselves from without; we see from within, as participants and therefore as partial beings. Neutrality, after all, is another position, not the absence of one. "There is no secure exteriority, only the restless movement between what is said and the act of saying it" (Demo; Friebel, 2025, p. 76). This caution also applies to informal assessment, because judgments formed in everyday interactions may either open learning opportunities or generate labels and exclusions; they must therefore be conducted ethically and encouragingly (Villas Boas, 2013).

Given the inevitably partial character of assessment, its contestability must be preserved: assessment is always open to question, and it is essential that it remain so, because this curbs evaluative arrogance and compels assessors to recognize the limits of their position. The issue, therefore, "is not merely to recognize the disciplinary character of teaching, but to consider how this process can be transformed into a path toward emancipation rather than blind obedience" (Demo;

Friebel, 2025, p. 99). In society, one of the most disputed forms of authority is the authority to determine what counts as valid, because authority ultimately consists in imposing validities. Paternal authority may be condensed into this function: determining what counts within the household. Aware of this, Habermas (1989; 2002) proposed that truth is a validity claim, both because every truth has an owner - no human being is wholly truthful, since evolution does not sustain such (im)purity - and because truth is contested. It cannot be separated from the dialectics of power. In an open dialectics of power, validity claims are disputed; in a linear, dictatorial context, they are imposed.

At the pedagogical level, dialogical assessment must be distinguished from autocratic assessment. In the former, the student remains a subject of the process and can respond, argue, question, and disagree; in the latter, the student is reduced to an object of judgment. Large-scale assessments such as Ideb do not fully permit this reciprocity and therefore should not constitute the sole mode of assessment.

From a mediating perspective, partial results should not function as final verdicts, but as information that helps interpret the learner's trajectory and reorganize pedagogical action (Hoffmann, 2012).

To avoid confusing “contestable” with “deficient,” it is useful to recover the dialectics of negation within critique. We often adopt a hasty view of what is contestable and emphasize only its deficient aspect. In open dialectics, which addresses relationships that remain open, formative, and developmental, deficiency need not be purely negative, not least because negation is constitutive of dialectics. When we claim that only constructive criticism matters, we commit a revealing contradiction: criticism is criticism only insofar as it negates; purely positive criticism criticizes nothing. Bachelard (1984; 1996; Demo, 2025) strongly emphasized “no” as an epistemological principle, because learning to think begins with deconstruction and proceeds through reconstruction while remaining open - modus tollens, in epistemological terms. For many authors, science cannot “verify” a theory, because no local induction can be generalized (Popper, 1959; Kolakowski, 1972), but it can negate or falsify a theory because a single negative case may suffice.

This remains controversial (Demo, 2011), but the questioning character of science should be preserved as one of its major contributions (Syed, 2019): knowing how to think means questioning, above all questioning oneself (Demo, 2010).

Once fallibility is recognized as intrinsic to assessment, the central problem becomes peremptoriness. A contestable assessment is, by its very nature, fallible and situated within evolutionary and sociohistorical dynamics. Peremptory assessments are authoritarian, miseducative, and deforming. “Scientific truth, when detached from self-criticism and ethical responsibility, becomes dogma” (Demo; Friebel, 2025, p. 51). It is true that we do not like being assessed, and this applies above all to assessors: they make a living by assessing others, yet detest being assessed themselves. Avoiding assessment, however, is counterproductive, both because in social life we assess and counter-assess one another, valuing and devaluing one another, and because we cannot step outside or exempt ourselves from the dialectics of power, in which we are deeply and naturally embedded (Foucault, 2000; Bourdieu, 2007).

In debates about classification, pedagogical judgment must be distinguished from hierarchical ranking. Every assessment entails interpreting evidence and formulating provisional judgments, but this does not require ranking students against one another. In mediating assessment, the focus falls on the student's singular trajectory and on the interventions needed to support progress, not on the student's position relative to others (Hoffmann, 2012). Temporal comparisons may be useful when students are monitored in relation to their own trajectories, without converting assessment into competition.

In selection processes such as university entrance examinations and the ENEM, classification serves a specific social function: ranking candidates when places are limited. That purpose, however, should not be confused with formative assessment conducted in everyday school life. At the heart of sociological critique, Bourdieu identifies “distinction” as a driving force of schooling. One of the most incisive - and bitter - contributors to this debate, Bourdieu develops this argument most

forcefully in Distinction (2007), one of his most sarcastic and caustic works: we do not attend school in order to become “equal,” which is not even feasible - we are naturally unequal, or equal and diverse (Demo; Calisto, 2025b; McMahon, 2023) - but to distinguish ourselves, secure a better place in the sun, and compete for opportunities. Pedagogy often seeks to conceal this struggle, a gesture that becomes incoherent when it invokes emancipation, which is itself a form of struggle (Freire, 1997).

When Assessment Does Not Produce Learning

Starting from the premise that pedagogical change requires an understanding of the concrete conditions of learning, diagnosis is understood here as a process-based, contextual, and revisable interpretation of educational trajectories. The purpose is neither to locate the sole cause of non-learning in the learner nor to fix the learner within a label, but to interpret evidence and pedagogical conditions in order to decide on new interventions. From this perspective, error is part of knowledge construction and can guide renewed teaching, feedback, and rewriting rather than operate as a definitive sign of incapacity (Luckesi, 1990; 2011). The available data are alarming: in 2023, among public upper-secondary schools, the proportion of students meeting the Adequate Learning benchmark in mathematics ranged from 2% to 9% across the states, with a national average of 5%. This situation suggests that teaching cannot be defended merely on the grounds that lessons were delivered but must be justified by an effective commitment to learning: “the teacher's pedestal is not the lesson, but the teacher's commitment to the student's learning, which must always precede the commitment to the teacher's own continuing learning” (Demo, as cited in Friebel, 2025, p. 52). A lesson may have been delivered, but when even a minimal learning effect fails to appear, the system reveals its lack of productivity.

The problem, of course, does not begin in upper-secondary education. It accumulates throughout the educational trajectory: mathematics already lags behind in the early years of elementary school, and this

inequality intensifies until learning becomes residual. To make the diagnosis visible, it is useful to begin with the figures for public upper-secondary education. Even when some results “improve,” the situation remains severe: lessons are delivered, but they fail to produce a minimally acceptable learning effect. In this context, remedial education programs, continuing teacher education focused on structural inequalities, and targeted strategies to improve mathematics teaching among vulnerable populations may be essential to reducing disparities and ensuring adequate learning for all (Demo; Friebel, 2025, pp. 144-145).

In Portuguese language, still within the public school system, the percentages are relatively higher, ranging from 19% to 43%, with a national average of 32%. Objectively, however, they remain insufficient and do not “rescue” public schooling as an institution of learning. The discrepancy between subject areas is revealing: even when the result in Portuguese language is several times higher than that in mathematics, this does not mean that an acceptable formative level has been achieved. It means, rather, that mathematics is in even more explicit collapse.

Turning to private schools, the comparison indicates that the problem is also systemic and instructionist. In 2023, among private upper-secondary schools, the proportion meeting the Adequate Learning benchmark in mathematics ranged from 15% to 49%, with a national average of 31%. This result requires qualification: it largely reflects the performance of the socioeconomic elite in a fee-paying school system, and yet the outcome remains disappointing, particularly when it is presented as “validation” of pedagogical quality. Performance in Portuguese language was better, ranging from 48% to 77%, with a national average of 67%, but it is still not fully satisfactory given what should be expected from schools with superior material conditions. The discrepancy between Portuguese language and mathematics is therefore not merely a “statistical detail”; it points to a recurrent pattern of unequal learning and a persistent weakness in mathematics education, even in privileged contexts.

Once it is acknowledged that private schools perform “better,” the question becomes why such an unproductive system continues to be maintained. The point is not to deny differences between school systems, but to reject the normalization of mediocre results as though they were acceptable. The unavoidable question is this: knowing all this, why do we continue to preserve such an ineffective system? Do we assess in order not to assess? If the ENEM exposes year after year that almost no one writes well, why do we persist with a system that does not make writing a sustained practice? Who benefits from such meticulously produced deprivation?

The availability of an indicator, however contestable, brings into view the critique of the Ideb Proficiency score (0-10), its most widely disseminated form because it condenses mathematics and Portuguese language into a single figure. The pedagogical validity of this formulation can be challenged. First, it makes little sense to calculate an “average” between curricular subjects that are not learned as an average but in their own integrity. Second, because performance may diverge sharply, particularly in upper-secondary education, the average tends to conceal the crisis in mathematics. To compare these results, it is necessary to distinguish the Ideb performance component, derived from standardized proficiency scores from the Basic Education Assessment System (Sistema de Avaliação da Educação Básica, Saeb), from the percentage of students classified as having achieved “Adequate Learning.” The former enters the calculation of Ideb together with the pass rate; the latter represents the proportion of students who reach a specified benchmark in each curricular subject. The two indicators should therefore not be treated as equivalent scales, although they may be analyzed complementarily. In 2023, for example, Adequate Learning in the public system was 5% in mathematics and 32% in Portuguese language - a substantial difference. A numerical average of these percentages explains little pedagogically and may politically conceal the critical point.

Adding to proficiency a component not covered by IRT creates a possible disjunction between rising pass rates and learning that has actually

been demonstrated. “If education has a possible horizon, it cannot be based on the illusion of numbers that disguise reality” (Demo; Friebe, 2025, p. 123). Because the pass rate is not covered by IRT and is vulnerable to institutional pressures, it may “improve” the final indicator without changing actual learning. In 2023, the Ideb score for upper-secondary education was 4.1, resulting from a proficiency average of 4.45 in Portuguese language and mathematics multiplied by a pass rate of 0.91. This calculation illustrates how passing functions as a factor that “sweetens” the diagnosis.

To clarify how the scale can soften the problem, it is useful to compare the Ideb Proficiency score with Adequate Learning. In public schools in 2023, the Ideb Proficiency score was 5.7 in the early years of elementary school, 4.7 in the later years, and 4.1 in upper-secondary education. The decline across stages appears systematic and continuous. When Adequate Learning is disaggregated by subject, however, the picture changes: the issue is not merely a “gradual decline,” but a widening chasm, particularly in mathematics. In Portuguese language, the percentages were 55% in the early elementary years, 36% in the later elementary years, and 32% in upper-secondary education. In mathematics, the corresponding figures were 44%, 16%, and 5%, indicating a dramatic rupture in the transition to the later elementary years and collapse in upper-secondary education.

This comparison supports the argument that Adequate Learning, although itself contestable, is more revealing and less lenient, particularly because it separates subject areas and expresses the severity of the problem in percentages. The preferential use of the Ideb Proficiency score in school culture and educational management may be contributing to a less unsettling diagnosis, as though “on average” progress were being made. Yet what the average conceals is central: mathematics has stagnated, and in upper-secondary education the result approaches a statistical residual. This suggests an avoidance of the severity of diagnosis: “a mechanism of illusion in which high numbers conceal increasingly shallow learning” (Demo; Friebe, 2025, p. 122).

At the level of school culture, the problem shifts to the overvaluation of teaching as the principal “event.” Education has a persistent habit of placing teaching - and the teacher's own centrality - above everything else. We still assume that the major event in the classroom is the lesson. It never was. The major event has always been student learning. When learning does not occur, teaching tends to become an empty performance.

With the advent of artificial intelligence (AI), empty teaching is likely to be overtaken, which makes authorship and a maieutic approach even more urgent. Lessons can, of course, be relevant when taught by teachers who are themselves authors; as a general rule, however, AI exposes instructionism because it can repackage content more effectively than we can. The revolution AI brings to education is syntactic rather than semantic, since AI is an automated tool (Setzer, 2021). The teacher's role therefore shifts toward a maieutic function: caring for and cultivating the student's authorship. Khan controversially tends to attribute this task to AI (Khan, 2025), whereas authorship in fact requires experience, intention, responsibility, and relationship.

Returning to the pass/fail dilemma, the criticism concerns passing students without learning, a problem now at risk of intensification. In education, we argue that failing students should be avoided - if a student has done what was possible, there is no reason to fail them - but students cannot be passed without learning. Yet schools routinely pass students who have not learned. Advancement, from this perspective, is not synonymous with bureaucratic passing; it means creating conditions for the learner to make qualitative progress (Hoffmann, 2017). AI may exacerbate the problem by offering simplified and “personalized” content as a product for immediate consumption, reinforcing the old hypostatization of content over learning. The value of learning lies in its singular, lived, and personal semantics, which requires full student agency and attentive teacher care. AI does not know this: an algorithm does not think, care, feel, or live.

To close this discussion, assessment must shift from content toward authorial learning, including when the use of AI is evaluated. It is not enough to assess what was answered; it is necessary to assess how learning occurred, how arguments were constructed, how texts were rewritten, and how authorship was assumed. AI is a technology - a means, not an end - and because it is here to stay, it must be approached critically: controlled before it controls us and educated before it diminishes us. We should neither embrace it gullibly nor resist it quixotically; we should diagnose and care.

Conclusion

The indicators examined in this analysis - from ENEM to Ideb, including evidence of longitudinal decline, as observed in the Federal District, and the contrast between public and private school systems - do more than confirm low performance. They reveal a persistent disjunction between schooling and learning, one that endures despite the repeated availability of evidence. The problem is not a lack of data, but the way assessment has been subsumed by an instructionist logic: measurement is institutionalized, while the diagnosis that would compel change is resisted. In this context, educational failure becomes normalized, and a school model centered on classroom instruction as an end in itself is perpetuated, operating as though teaching alone were sufficient even when learning does not occur.

Throughout the essay, the assessment of learning emerges as a qualitative challenge that cannot be addressed by remote external assessments or standardized tests that confine learning to what is measurable and exclude what matters most: authorship, elaboration, progression, and rewriting, supported by close teacher guidance and process-based records such as digital portfolios. The ENEM essay, despite its formulaic format, inadvertently signals what is structurally absent: writing remains exceptional in school when it should be an everyday practice, because deep learning requires sustained engagement in authorial work.

The analysis also shows that assessment can be neither “pure” nor wholly fair: as a human act, it is inevitably partial and embedded in relations of power. The problem, however, lies not in acknowledging this fallibility, but in turning it into finality - as though the evaluator could occupy an untouchable, elevated position. Preserving the contestable character of assessment curbs evaluative arrogance and opens space for a dialogical practice in which students remain subjects of the process and are able to respond, argue, disagree, and learn - the opposite of what occurs when assessment becomes a disciplinary device, a mechanism of distinction, and a form of brutal social selection. If schooling is a contested terrain, the pedagogical task is not to conceal this conflict but to regulate it, directing it toward greater equality of opportunity and a cooperative culture that does not allow competition to undermine learning.

Finally, the analysis suggests that the principal driver of stagnation is the avoidance of diagnosis: indicators are softened, pass rates are manufactured, and an appearance of normality is produced. The critique of Ideb proficiency scores and of pass rates as cosmetic devices illustrates this mechanism: by downplaying the severity of the problem, the system, its classroom routines, and its institutional practices are preserved, even as the figures - especially in mathematics - approach collapse. In this essay, diagnosis does not refer to clinical diagnosis. Rather, it denotes a pedagogical, contextual, and revisable mapping of learning, connected to ongoing monitoring and decision-making. Making this boundary explicit helps prevent difficulties, differences, and errors from being recast as symptoms or individual deficits, a risk identified in critical scholarship on the medicalization of education (Meira, 2012; Scarin; Souza, 2020). Without such pedagogical interpretation, care may be reduced to improvisation; when it guides feedback, rewriting, and intervention, assessment effectively fulfills its formative function.

It is therefore unproductive to sustain the myth that one day we will devise an assessment that renders itself unnecessary, as though a

depoliticized school, free of conflict and cured of the distortions of power, were possible. The decisive question is different: how to design and sustain assessment as a practice of care. The arrival of artificial intelligence makes this imperative even more explicit: checking products and reproducing content are no longer sufficient, because AI performs such tasks efficiently and continuously. If there is a possible gain, it lies in restoring to schools a fundamental insight: the central event of the classroom has never been instruction itself, but learning.

The specific contribution of this essay lies in conceptualizing care not merely as an affective disposition or a welcoming attitude, but as a pedagogical principle that guides assessment practice. In this context, caring means continuously following learning trajectories, interpreting difficulties and potentialities, providing high-quality feedback, and ensuring opportunities for revision, rewriting, and reconstruction. This conception also requires rejecting pedagogical abandonment when learning does not occur: assessment results must lead to intervention, the reorganization of teaching, and renewed opportunities to learn. Care-oriented assessment therefore neither lowers expectations nor conceals difficulties; rather, it transforms the evidence produced through assessment into a concrete commitment to learning and to the learner's authorship.

References

BACHELARD, Gaston. *A filosofia do não: filosofia do novo espírito científico*. 3. ed. Lisboa: Editorial Presença, 1984.

BACHELARD, Gaston. *O novo espírito científico*. Lisboa: Edições 70, 1996.

BOURDIEU, Pierre. *A distinção: crítica social do julgamento*. São Paulo: Edusp; Porto Alegre: Zouk, 2007.

CASEMIRO, Poliana; SANTOS, Emily. Enem 2024: apenas 12 candidatos conseguem nota máxima e número é o menor em dez anos. *G1 Educação*, 13 jan. 2025. Disponível em: <https://g1.globo.com/educacao/enem/2024/noticia/2025/01/13/enem-2024-coletiva-mec-inep.ghtml>. Acesso em: 14 jul. 2026.

DEHAENE, Stanislas. *É assim que aprendemos: por que o cérebro funciona melhor do que qualquer máquina (ainda...)*. São Paulo: Contexto, 2022.

DEMO, Pedro. *Pesquisa e informação qualitativa: aportes metodológicos*. Campinas: Papirus, 2001.

DEMO, Pedro. *Saber pensar é questionar*. Brasília, DF: Liber Livro, 2010.

DEMO, Pedro. *A força sem força do melhor argumento: ensaio sobre 'novas epistemologias virtuais'*. Brasília, DF: IBICT, 2011.

DEMO, Pedro. *Aprender como autor*. São Paulo: Atlas, 2015.

DEMO, Pedro. *Atividades de aprendizagem: sair da mania do ensino para comprometer-se com a aprendizagem do estudante*. Campo Grande: SED/MS, 2018. Disponível em: <https://drive.google.com/file/d/1FKskDCxNB422PVhrjrDjD48S4cjsb77-/view>. Acesso em: 14 jul. 2026.

DEMO, Pedro. A filosofia do não. *Blog do Pedro Demo*, 2025. Disponível em: <https://pedrodemo.blogspot.com/2025/09/ensaio-1326-filosofia-do-nao.html>. Acesso em: 14 jul. 2026.

DEMO, Pedro; CALISTO, Cristiano de Souza. *Pesquisa qualitativa precisa de fundamentos epistemológicos: argumentando pela necessidade da pesquisa qualitativa para dar conta de realidades complexas, sobretudo do mundo da vida*. Jundiaí: Paco Editorial, 2025.

DEMO, Pedro; CALISTO, Cristiano de Souza. A quem interessa uma escola tão inepta? *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 33, n. 128, p. 1-23, 2025a. DOI: <http://doi.org/10.1590/S0104-4036202500330491>.

DEMO, Pedro; CALISTO, Cristiano de Souza. *Direitos humanos iguais e diversos: desafios do igualitarismo*. Jundiaí: Paco Editorial, 2025b.

DEMO, Pedro; FRIEBEL, Silvane. *Ensino morreu. Amém!* Jundiaí: Paco Editorial, 2025.

FOUCAULT, Michel. *A ordem do discurso*. São Paulo: Loyola, 2000.

FREIRE, Paulo. *Pedagogia da autonomia: saberes necessários à prática educativa*. Rio de Janeiro: Paz e Terra, 1997.

FRIGOTTO, Gaudêncio. *A produtividade da escola improdutiva: um (re)exame das relações entre educação e estrutura econômico-social capitalista*. São Paulo: Cortez, 1988.

FRIGOTTO, Gaudêncio. A produtividade da escola improdutiva 30 anos depois: regressão social e hegemonia às avessas. *Revista Trabalho Necessário*, [S. l.], v. 13, n. 20, p. 206-233, 2015. Disponível em: <https://periodicos.uff.br/trabalhonecessario/article/view/8619>. Acesso em: 14 jul. 2026. DOI: <http://doi.org/10.22409/tn.13i20.p8619>.

FRIEBEL, Silvane. *Alfabetização e autoria pela intervenção pedagógica em escola pública do Distrito Federal no contexto dos direitos humanos e cidadania*. Curitiba: CRV, 2025.

HABERMAS, Jürgen. *Consciência moral e agir comunicativo*. Rio de Janeiro: Tempo Brasileiro, 1989.

HABERMAS, Jürgen. *Agir comunicativo e razão descentralizada*. Rio de Janeiro: Tempo Brasileiro, 2002.

HATTIE, John; TIMPERLEY, Helen. The power of feedback. *Review of Educational Research*, [S. l.], v. 77, n. 1, p. 81-112, 2007. DOI: <http://doi.org/10.3102/003465430298487>.

HOFFMAN, Donald D. *The case against reality: why evolution hid the truth from our eyes*. New York: W. W. Norton & Company, 2019.

HOFFMANN, Jussara. *Avaliação mediadora: uma prática em construção da pré-escola à universidade*. 32. ed. Porto Alegre: Mediação, 2012.

HOFFMANN, Jussara. *Avaliar para promover: as setas do caminho*. 16. ed. Porto Alegre: Mediação, 2017.

INAF – INDICADOR DE ALFABETISMO FUNCIONAL. *Inaf 2024: legado de futuro do alfabetismo funcional*. São Paulo: Ação Educativa, 2025. Disponível em: <https://alfabetismofuncional.org.br/dados-inaf-2024.pdf>. Acesso em: 14 jul. 2026.

KHAN, Salman. *Ensinar tudo a todos: como a IA vai revolucionar a educação*. [S. l.]: HarperCollins Brasil, 2025.

KOLAKOWSKI, Leszek. *Positivist philosophy: from Hume to the Vienna Circle*. London: Penguin Books, 1972.

LIBÂNEO, José Carlos. O dualismo perverso da escola pública brasileira: escola do conhecimento para os ricos, escola do acolhimento social para os pobres. *Educação e Pesquisa*, São Paulo, v. 38, n. 1, p. 13-28, jan./mar. 2012.

LUCKESI, Cipriano Carlos. Prática escolar: do erro como fonte de castigo ao erro como fonte de virtude. In: FUNDAÇÃO PARA O DESENVOLVIMENTO DA EDUCAÇÃO (org.). *Caderno Ideias*, n. 8. São Paulo: FDE, 1990. p. 133-140.

LUCKESI, Cipriano Carlos. *Avaliação da aprendizagem escolar: estudos e proposições*. 22. ed. São Paulo: Cortez, 2011.

MCMAHON, Darrin M. *Equality: the history of an elusive idea*. New York: Basic Books, 2023.

MEIRA, Marisa Eugênia Melillo. Para uma crítica da medicalização na educação. *Psicologia Escolar e Educacional*, São Paulo, v. 16, n. 1, p. 135-142, jan./jun. 2012. DOI: <http://doi.org/10.1590/S1413-85572012000100014>.

PIAGET, Jean. *A construção do real na criança*. São Paulo: Ática, 1996.

POPPER, Karl R. *The logic of scientific discovery*. London: Hutchinson, 1959.

QEDU. *Dados educacionais do Brasil: Saeb e Ideb 2023*. [S. l.], 2023. Disponível em: <https://qedu.org.br/brasil>. Acesso em: 14 jul. 2026.

ROITMAN, Isaac. A crise da educação no Brasil não é uma crise: é projeto. *UnB Notícias*, Brasília, DF, 14 out. 2022. Disponível em: <https://noticias.unb.br/artigos-main/6081-a-crise-da-educacao-no-brasil-nao-e-uma-crise-e-projeto>. Acesso em: 14 jul. 2026.

SCARIN, Ana Carla Cividanes Furlan; SOUZA, Marilene Proença Rebello de. Medicalização e patologização da educação: desafios à psicologia escolar e educacional. *Psicologia Escolar e Educacional*, [S. l.], v. 24, e214158, 2020. DOI: <http://doi.org/10.1590/2175-35392020214158>.

SETZER, Valdemar W. *IA: inteligência artificial ou imbecilidade automática? As máquinas podem pensar e sentir?* 2. ed. São Paulo: Biblioteca 24 Horas, 2021.

SYED, Matthew. *Rebel ideas: the power of diverse thinking*. London: John Murray, 2019.

VAROUFAKIS, Yanis. *Technofeudalism: what killed capitalism*. [S. l.]: Vintage Digital, 2023.

VILLAS BOAS, Benigna Maria de Freitas. *Portfólio, avaliação e trabalho pedagógico*. Campinas: Papirus, 2004.

VILLAS BOAS, Benigna Maria de Freitas. Avaliação formativa e formação de professores: ainda um desafio. *Linhas Críticas*, Brasília, DF, v. 12, n. 22, p. 75-90, jan./jun. 2006. DOI: <http://doi.org/10.26512/lc.v12i22.3283>.

VILLAS BOAS, Benigna Maria de Freitas. Avaliação da educação básica: das informações existentes ao interior das escolas. *Retratos da Escola*, [S. l.], v. 7, n. 12, p. 167-178, 2013. DOI: <http://doi.org/10.22420/rde.v7i12.271>.

VILLAS BOAS, Benigna Maria de Freitas; DIAS, Elisângela Teixeira Gomes. Provinha Brasil e avaliação formativa: um diálogo possível? *Educar em Revista*, Curitiba, n. esp. 1, p. 35-53, 2015. DOI: <http://doi.org/10.1590/0104-4060.41421>.

Received in January 2026.

Approved in may 2026.