DOSSIER

Quality, learning and systemic assessment: discourses from international organizations for Latin American countries

Assessment methods for learning: what do teachers from Latin American countries show in the PISA 2018 contextual questionnaire?

Métodos de avaliação para aprendizagem: o que evidenciam os professores de países latino-americanos no questionário contextual do PISA 2018?

Ronildo Stieg[a]
ronildo.stieg@yahoo.com.br

Denilson Junio Marques Soares[b]
denilson.marques@ifmg.edu.br

Wagner dos Santos[a]
wargnercefd@gmail.com

ABSTRACT

The Program for International Student Assessment (PISA) is a standardized exam carried out every three years by the Organization for Economic Co-operation and Development to measure the educational level of 15-year-old students in reading, mathematics, and science. This study aims to analyze what and how often teachers use methods to assess their students' learning other than PISA. This mixed research uses the responses of teachers from five Latin American countries (Brazil, Chile, Panama, Peru, and the Dominican Republic) who responded to the 2018 PISA contextual questionnaire. Results organized the discussions based on three categories of analysis that align with assessment methods for learning, focusing on the dimensions of a) hetero-assessment, b) self-assessment, and c) feedback. Predominantly, hetero-assessment was the central approach, reflecting a traditional emphasis on the active role of the evaluating teacher. The absence of co-assessment and shared assessment highlights a gap in assessment practices, indicating the need to promote more collaborative and peer-to-peer assessments. This study concluded that these results must be interpreted considering contextual differences and local policies, showing the complexity of evaluation practices and indicating the importance of adopting comprehensive and collaborative methods to improve more inclusive educational environments that concern themselves with the learning processes of students in basic education.

Keywords: System Evaluation. PISA. Educational Policies. Learning. Latin America.

ª Universidade Federal do Espírito Santo (UFES), Vitória, Espírito Santo, Brazil.

ᵇ Instituto Federal de Minas Gerais (IFMG), Ouro Preto, Minas Gerais, Brazil.

RESUMO

O Programa de Avaliação Internacional de Estudantes (PISA) é um exame padronizado realizado a cada três anos pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), com o objetivo de mensurar o nível educacional de jovens de 15 anos em Leitura, Matemática e Ciência. O objetivo deste artigo consiste em analisar quais são e com que frequência os docentes utilizam métodos de avaliação da aprendizagem de seus alunos para além do PISA. Caracteriza-se como uma pesquisa de natureza mista que utiliza as respostas dos professores de cinco países da América Latina (Brasil, Chile, Panamá, Peru e República Dominicana) que responderam ao questionário contextual do PISA de 2018. Os resultados permitiram organizar as discussões a partir de três categorias de análises que se alinham com os métodos de avaliação para aprendizagem, concentrando-se nas dimensões: a) heteroavaliação; b) autoavaliação; e c) retroalimentação. Predominantemente, a heteroavaliação foi a abordagem central, refletindo uma tradicional ênfase no papel ativo do professor avaliador. A ausência de coavaliação e avaliação compartilhada destaca uma lacuna nas práticas avaliativas, indicando a necessidade de promover avaliações mais colaborativas e entre pares. Conclui que esses resultados devem ser interpretados considerando as diferenças contextuais e políticas locais, revelando a complexidade das práticas avaliativas e indicando a importância de adotar métodos abrangentes e colaborativos para aprimorar ambientes educacionais mais inclusivos e preocupados com os processos de aprendizagem dos estudantes da educação básica.

Palavras-chave: Avaliação de Sistema. PISA. Políticas Educacionais. Aprendizagem. América Latina.

Introduction

Large-scale assessments, present in different contexts and with varied (local, national, and international) scopes, have configured mechanisms to formulate public policies for socioeconomic development and education quality. Of these, the Programme for International Student Assessment (PISA) plays a notably prominent role in this scenario.

For Cotta (2001), its approach requires teachers to develop effective pedagogical strategies to promote good-quality education in its participating countries. The strategies the author highlights include a) the development of 21st-century skills, including critical thinking and collaboration; b) the interdisciplinary integration of real content and contexts; c) the promotion of reading and textual comprehension; d) the diversification of teaching methods with active learning and technology; e) the implementation of regular formative assessments; f) the promotion of inclusion and diversity, ensuring equitable opportunities; g) the stimulation of metacognition to promote autonomy; h) the facilitation of collaboration between teachers; i) the focus on languages and literacy, which are essential for academic success; and j) openness to innovation, integrating new technologies into the teaching and learning processes.

Thus, PISA, recognized for its international scope, offers results on advances and deficiencies in teaching, guiding national education plans and influencing changes in the education system, especially in the pedagogical practices of teachers, including their assessments (Schleicher, 2019; Jolandek; Pereira; Mendes, 2019; Fernandes *et al.*, 2018).

We highlight that PISA is applied to 15-year-old students regardless of grade/school year. In total, 79 countries participated in the 2018 initiative of the Organization for Economic Cooperation

and Development, including Brazil and Portugal (Schleicher, 2019). The data of this assessment — based on the strengths and weaknesses of the educational systems that have adopted it (Carnoy *et al.*, 2015; Soares; Candian, 2007) — encompasses reading, mathematics, and sciences and can be used by teachers but are more commonly used by educational management bodies and companies hired by education systems to promote the formulation of educational policies

The academic performance of students who participate in PISA assessments, measured by grades, is directly related to students' knowledge — especially mastery of reading, an indicator of the assessed skills (Esteban, 2000; Sassaki *et al.*, 2018). Factors such as the availability of books, family background (including parents' income and education), and personal characteristics have positively influenced the performance and the development of a taste for reading (Miranda *et al.*, 2015; Pinto; Bixirão Neto; Carvalho, 2019). Moreover, elements of the educational system, teacher training, cultural context, and educational practices significantly influence reading performance (Salokangas; Kauko, 2015; Bartholomeu *et al.*, 2016; Matos; Ferrão, 2016).

As Schleicher (2019) emphasizes, PISA only serves as a strategic tool for changes in the education system, functioning as a predictive indicator of students' mathematics, reading, and science abilities. Family background, home and school environment, and socioeconomic factors play significant roles in academic performance, evincing positive correlations with socio-emotional skills (Fernandes *et al.*, 2018; Sassaki *et al.*, 2018; Miranda *et al.*, 2015). These findings highlight the relevance of PISA as a comprehensive indicator to guide the formulation of public policies and decision-making by educational managers (Carvalho, 2016).

In addition to being correlated with the mastery of content and deep understanding of the concepts of certain areas, PISA results have supported a narrative about the training (or its absence) in education. However, the uses of PISA focused on accountability and in the creation of rankings have limited its ability to effectively guide educational policies. Instead of promoting informed decisions, such as the development of educational plans that meet various contextual realities, these uses often result in subjective evaluations that fail to necessarily be useful for making concrete decisions (Verger; Fontdevila; Parcerisa, 2019).

In view of this situation, we raise the following questions: what methods of learning assessment do teachers in different Latin American countries use? To what extent has PISA made it possible to understand the impacts of this evaluation on the assessment practices of teachers in their daily school lives? Is it possible to identify similarities between the evaluation methods of the teachers who responded to the 2018 PISA contextual questionnaire?

The relevance of this research stems from its aim to support discussions on the policy of a large-scale assessment in education (PISA) across countries, providing inputs on the evaluation methodologies teachers use in their pedagogical practice and that can influence the results of PISA and its uses in educational policies. The collected information may guide the path of educational policies in different countries, considering their international and national influences, especially discussing the extent to which teachers have used methods to assess student learning according to the 2018 PISA questionnaire.

Thus, this study aims to analyze which methods of assessing student learning are frequently used by teachers in five Latin American countries (Brazil, Chile, Panama, Peru, and the Dominican Republic). For this, it used teachers' answers to the 2018 PISA contextual questionnaire. This study hypothesized that the PISA questionnaire fails to evaluate co- and shared assessment, although it proposes to broadly assess learning.

Theoretical Framework

This study delimits some of the widely disseminated elements in academia as its theoretical-conceptual framework, including the definition of large-scale assessments in education according to Ramos-Zincke (2018) and Contreras and Torres (2023), who seek to define the quality of teaching and learning and, in some cases, to outline profiles of students and society in line with the neoliberal policies in force in education.

Given the adoption of these characteristics, large-scale assessments face substantial challenges regarding their purposes (Vanlommel; Schildkamp, 2019), often fitting into what is internationally known as accountability systems. These systems focus on assessing student performance in relation to specific curriculum content (Ozga, 2020), which this study takes as PISA.

A second concept this study mobilizes corresponds to the broadly interpreted assessment for learning. It refers to obtaining evidence related to students' specific situation to provide feedback that will help them move forward. It emphasizes students' self-regulation and autonomy regarding their learning, a competence they develop by self-assessment and dialogue in the classroom (Black; Willim, 2006).

Another concept that involves evaluation practices refers to the centrality of those who perform the assessment, understood as its agents, which includes self-assessment, hetero-assessment, and co-assessment (Castejón Oliva *et al*., 2011; Buscà *et al*., 2010). Self-assessment occurs when individuals or groups evaluate their own actions and learning processes (Casanova, 1997). Represented by the logic of subjectivity (to the extent that the self-evaluated subject becomes part of the evaluations and becomes aware of their successes and failures), self-assessment is attributed the idea of self-learning and/or personal autonomy to interrelated elements, leading to the development of an educational system with permanent formative attitudes (Castillo; Cabrerizo, 2012).

Hetero-assessment is one in which the evaluator and the evaluated are not the same person and is usually carried out within the educational center (for example, when teachers evaluate their students or when management teams evaluate some aspects of schools) (Casanova, 1997; Castillo; Cabrerizo, 2012). Marked by teachers' follow-up, monitoring, and guidance, it imposes external control, evaluating and judging students' learning and teachers' own teaching and academic activity in general.

The third agent of evaluation, co-assessment, occurs when people or groups belonging to an educational center evaluate each other, alternating their roles (Castillo; Cabrerizo, 2012), sharing responsibility and enabling the continuous intercommunication of the entire teaching and learning process (Stieg *et al*., 2022).

Although feedback is sometimes used as a retroregulation, feedback, or reference, has its application in education as it determines that the achieved results influence the initial forecast and each phase of the assessment (Stieg, 2022). Feedback, exercised from the teaching-learning processes, helps to redirect the elements that make up the didactic model, overcoming the instrumental and referential character of evaluation (Cardona, 1994) and favoring the desire to correct the aspects that can be improved in the dynamics of teaching and learning (Gosálbez, 1989).

Thus, analyzing the sources from this theoretical framework studies of large-scale learning assessments by teachers' responses to the PISA 2018 contextual questionnaire. These answers can offer clues (Ginzburg, 1989) to how assessment for learning, assessment agents, and feedback feature the assessment methods used by teachers in five Latin American countries.

Methodology

This mixed (quantitative and qualitative) (Creswell; Plano Clark, 2013) research is based on the comparative method (Bloch, 1998). Its sources comprise teachers' answers to the following question included in the contextual questionnaire of the PISA 2018: "How often do you use the following methods of assessing student learning?," which offers seven response items: a) I develop and administer my own assessment; b) I administer a standardized test; c) I collect data from classroom assignments or home work; d) I have individual students answer questions in front of the class; e) I let students judge their own progress; f) I provide written feedback on student work in addition to a <mark, i.e. numeric score or letter grade>; and g) I observe students when working on particular tasks and provide immediate feedback.

These items were organized on a scale that offers participants four answer alternatives: "never or almost never," "some lessons," "many lessons," and "every lesson or almost every lesson." These options were initially coded as 1, 2, 3, and 4, respectively, for quantification. In this context, a higher average would signal a more regular frequency in the adoption of the mentioned practices.

Based on the intrinsic nature of the items, three categories of analysis were then established that align with the assessment methods for learning, focusing on the dimensions of hetero-assessment (Category 1), self-assessment (Category 2), and feedback (Category 3). The central objective of this categorization was to operationalize a comparative analysis using PISA data as a basis for these categories.

The countries selected for this analysis participated in the survey and are located in Latin America: Brazil (n = 8,969), Chile (n = 3,755), Panama (n = 3,632), Peru (n = 5,146), and the Dominican Republic (n = 2,700). The collected information was organized into tables and analyzed by descriptive measures, including arithmetic means, standard deviations, and coefficients of variation. The latter represents the division of the second by the first in percentages, offering a dimension on answer regularity. In this case, the lower the coefficient of variation, the greater the similarity between teachers' opinions. Gomes (1985) classifies the coefficient of variation as low if below 10%; as medium, if from 10 to 20%; as high, if from 20 to 30%; and as very high, if above 30%.

As the used scale is ordinal, Mann and Whitney's (1947) non-parametric U test was used to compare participants' responses. In summary, the difference between the answers was considered statistically significant when its p-value lied below 0.05 (5%), as per the standard in the literature. All analyses were performed on R, version 4.1.0 (R Core Team, 2021).

Results

The first category corresponds to the Assessment Methods for Learning that focus on hetero-assessment, which describes and discusses the answers to items A, B, and C (Table 1). The second category addresses the Assessment Methods for Learning that focus on self-assessment, which are addressed in items D and E (Table 2). The third category, based on the answers to items F and G (Table 3), discusses the Assessment Methods for Learning that focus on feedback.

Assessment methods for learning with a focus on hetero-assessment

Table 1 describes the descriptive and inferential statistics related to the first category of evaluation method in the PISA 2018 contextual questionnaire applied to teachers: hetero-assessment.

Table 1: Frequency with which teachers use learning assessment methods with a focus on hetero-assessment

Item A – I develop and administer my own assessment					
	Respondents	Mean	Comparison	Standard deviation	CV	CV classification
Brazil	5,212	3.11838066	a	0.758927315	24.33722493	High
Chile	2,367	2.932826362	b	0.90492699	30.85511646	Very High
Panama	1,945	3.63907455	c	0.655853742	18.02254207	Medium
Peru	3,692	3.137053088	a	0.788708769	25.14170932	High
Dominican Republic	1,899	2.883622959	b	0.937566757	32.51350021	Medium
Item B – I administer a standardized test					
	Respondents	Mean	Comparison	Standard deviation	CV	CV classification
Brazil	5,167	2.2663054	a	0.889214971	39.23632584	Very High
Chile	2,358	2.190839695	b	0.90754124	41.42435623	Very High
Panama	1,928	2.71473029	c	1.072840313	39.51922285	Very High
Peru	3,692	2.442307692	d	0.88912893	36.40527902	Very High
Dominican Republic	1,865	2.319571046	a	0.979559363	42.23019444	Very High
Item C – I collect data from classroom assignments or home work					
	Respondents	Mean	Comparison	Standard deviation	CV	CV classification
Brazil	5,206	2.80868229	a	0.874736946	31.14403326	Very High
Chile	2,372	2.756323777	b	0.889121804	32.25752398	Very High
Panama	1,937	3.453278265	c	0.718521715	20.80694517	High
Peru	3,693	2.992959653	d	0.795638148	26.58365765	High
Dominican Republic	1,910	3.452356021	c	0.731191525	21.17949367	High

Source: Prepared by the authors based on the data available in the PISA 2018 contextual questionnaire.
*Means followed by the same letter in the column fail to differ by Mann-Whitney test (p-value > 0.05).

Teachers' answers to items A, B, and C showed varied means of the assessments these professionals used between countries. Panama showed the highest averages (Table 1). However, it is important to note that item C showed no statistically significant difference between the means of the responses in Panama and the Dominican Republic. On the other hand, although Chile had the lowest averages, it showed no statistically significant differences with the Dominican Republic for item A.

Response variability showed a notable heterogeneity, especially in item B, the coefficient of variation of which was classified as very high in all analyzed countries. This scenario indicates no convergence of opinions among teachers in the different countries.

Analyses between countries show answer variability and point to differences in the methods of developing and applying assessments by teachers. This can suffer the influence from teaching methods, curriculum guidelines, or specific pedagogical approaches. On the other hand, the lower variability in Panama indicates an inverse movement to the previous case.

Assessment methods for learning by self-assessment

Table 2 shows the descriptive and inferential statistics related to the second category of evaluation method in the PISA 2018 contextual questionnaire applied to teachers: self-assessment.

Table 2: Frequency with which teachers use assessment methods for learning with a focus on self-assessment

Item D - I have individual students answer questions in front of the class						
	Respondents	Mean	Comparison	Standard deviation	CV	CV classification
Brazil	5,168	1.964009288	a	0.857640236	43.66782993	Very High
Chile	2,354	2.242141037	b	0.911546719	40.65519092	Very High
Panama	1,938	2.628998968	c	0.928354587	35.31209399	Very High
Peru	3,693	2.909017059	d	0.798629637	27.45359069	High
Dominican Republic	1,910	3.276963351	e	0.796749413	24.3136504	High
Item E – I let students judge their own progress						
	Respondents	Mean	Comparison	Standard deviation	CV	CV classification
Brazil	5,186	2.173158504	a	0.868582595	39.96867202	Very High
Chile	2,368	2.403293919	b	0.814214487	33.87910571	Very High
Panama	1,936	2.606404959	c	0.927673871	35.59208511	Very High
Peru	3,693	2.723260222	d	0.809698179	29.73267751	High
Dominican Republic	1,909	3.085385018	e	0.819191696	26.55071219	High

Source: Prepared by the authors based on the data available in the PISA 2018 contextual questionnaire.
* Means followed by the same letter in the column fail to differ by Mann-Whitney test (p-value > 0.05).

Teachers' answers to items D and E show an evident variation in the means of the assessments these professionals used in all countries, being more expressive in the Dominican Republic for both

items. On the other hand, Brazil registers the lowest averages, indicating a lower frequency in the adoption of the aforementioned practices. Thus, while Brazil shows a lower inclination for both self-assessment practices, Peru and the Dominican Republic seem to adopt these practices more often.

In turn, the coefficient of variation of the responses considerably varies for all countries, especially Brazil, Chile, and Panama, the classification of which — proposed by Gomes (1985) — is considered very high. This suggests that the use of methods that favor self-assessment vary more between teachers in these countries.

Assessment methods for learning with a focus on feedback

Table 3 shows the descriptive and inferential statistics related to the third category of evaluation method in the PISA 2018 contextual questionnaire applied to teachers: feedback.

Table 3: Frequency with which teachers make use of learning assessment methods with a focus on feedback

Item F – I provide written feedback on student work in addition to a <mark, i.e. numeric score or letter grade>						
	Respondents	Mean	Comparison	Standard deviation	CV	CV classification
Brazil	5,191	2.566364862	a	0.855691035	33.34253239	Very High
Chile	2,365	2.349682875	b	0.890592663	37.90267495	Very High
Panama	1,942	3.053553038	c	0.877229537	28.72815785	High
Peru	3,694	2.785056849	d	0.84815401	30.45374137	Very High
Dominican Republic	1,910	3.392146597	e	0.734804956	21.66194578	High
Item G – I observe students when working on particular tasks and provide immediate feedback						
	Respondents	Mean	Comparison	Standard deviation	CV	CV classification
Brazil	5,196	2.780408006	a	0.879853802	31.64477299	Very High
Chile	2,379	3.109289617	b	0.766464152	24.6507803	High
Panama	1,943	3.476582604	c	0.681603159	19.60555052	Medium
Peru	3,694	3.240119112	d	0.723571413	22.33163004	High
Dominican Republic	1,911	3.510727368	e	0.695176658	19.80149937	Medium

Source: Prepared by the authors based on the data available in the PISA 2018 contextual questionnaire.
* Means followed by the same letter in the column fail to differ by Mann-Whitney test (p-value > 0.05).

In both cases, we highlight that the Dominican Republic has the highest average, whereas Chile and Brazil, the lowest ones, respectively. It is important to note that these methods, which include the offer of written feedback on students' work, followed by a grade expressed in numbers or letters, may lead students to answer questions before their classes and enable them to evaluate their own progress. Table 3 shows that the observation of students during specific activities and the offer of feedback are considered practices.

The coefficient of variation in item G showed significant variability in the data, except for Panama and the Dominican Republic, which expressed a mean variability, suggesting a greater convergence in the responses of teachers in these two countries for these specific practices. In

summary, the teachers from the five countries indicate diverse preferences and approaches when providing feedback to students during the evaluation processes.

Discussion

The discourses of international organizations to Latin American countries on quality, learning, and systemic evaluation, must consider the intersection with assessment methods for learning, as evinced by the PISA 2018 contextual questionnaire. The analysis of teachers' evaluation practices in five Latin American countries – Brazil, Chile, Panama, Peru, and the Dominican Republic – shows a complex educational scenario in which the methods used by teachers to evaluate student progress vary significantly.

Focusing on the analysis of the data corresponding to the answers of the teachers who participated in the PISA contextual questionnaire in 2018 regarding evaluation methods showed three central aspects: methods that center assessment practices in the action of teachers (hetero-assessment); assessment practices that recognize the need to place students at the center of the assessment processes, demarcating autonomy and collaborative processes (self-assessment); and the assessment practices that function as feedback mechanisms for the teaching and learning processes (feedback).

The hetero-assessment category showed a notable disparity in evaluation means of between the analyzed countries. Its variability indicates that teachers use different methods to develop their own assessments, pointing to the influence of teaching methods, curricular guidelines, and specific pedagogical approaches. This first category evinces that the PISA questionnaire directs the items on the assessment methods used by teachers, reinforcing the idea of teacher-centered assessments within a summative evaluation function.

A summative assessment aims to find the level of learning students achieved, determine the effectiveness of all elements of the educational process, or even to classify and qualify by interindividual comparisons (Casanova, 1997). This form of assessment can find variables; specify measurable units or ranges; build evaluation instruments; and collect, elaborate, analyze, interpret, and practically apply data (Castillo; Cabrerizo, 2012). Perhaps because it is the function that predominates in the logic of the standardized exams, it still configures a preponderant option in the assessment practices of the teachers in this study.

The heterogeneity in the answers in the first category suggests no convergence of opinions between teachers across contexts, reflecting a varied perception of evaluation centered on teachers' action, hence the notable discrepancy between the averages of evaluations focused on hetero-assessment in all countries. This indicates that teachers have a different perception or use different methods to develop their own assessments when compared to the standardized test, similar to PISA, which suggests that assessment methods vary substantially.

Tamayo (2018) defines hetero-assessment as a type of evaluation in which the evaluator play a different role from the person who is evaluated. In the framework of education, hetero-assessments conventionally occur when teachers plan, design, and apply student evaluations. Generally, this type of evaluation is the most used in education, in which teachers may always have an active role.

According to Vera *et al.*, (2018), the teaching practice focused on hetero-assessment tends to center evaluations on the principles of hierarchy of teachers before students. On the other hand, these methods play an important role as long as different means and instruments are used.

This technique contrasts with peer evaluation (co-evaluation), in which students with the same condition score their performance mutually. It is a measure that should be objective and avoid involving personal sessions when evaluating and qualifying other individuals (Rodríguez; Ibarra; García, 2013). Thus, when performed correctly, results can be very beneficial.

Still establishing a comparative analysis between the answers of teachers from different countries, we found a similarity in the answers in item B, "I administer a standardized test." In this case, all countries had a high agreement rate (very high CV). On the other hand, when we compare these data with those presented in item A, "I develop and administer my own assessments," we find a significant disparity in CV between countries. These data offer elements to understand that standardized exams, such as PISA, have influenced the evaluation practice of teachers. All countries show a very high use of these tests. At the same time, the elaboration of evaluations by teachers has received less emphasis in hetero-assessment methods, except for Chile.

Assessment by standardized tests may indicate a concern of teachers to adapt their practices to external requirements since international, national, and local evaluations have occupied a central place in the educational public policies of these countries, reinforcing the concept of accountability and the production of rankings (Vanlommel; Schildkamp, 2019; Ozga, 2020). The concern from this process lies in reducing teacher autonomy associated with an externally oriented assessment that fails to primarily consider students' learning, previously defining the competencies and skills they should possess (Verger; Fontdevila; Parcerisa, 2019).

In the self-assessment category, assessment averages also vary significantly between countries. The Dominican Republic stands out with the most expressive averages, whereas Brazil shows the lowest ones. The considerable variability in the responses of all countries indicates that the use of methods that favor self-assessment is more variable among teachers. We note a lower inclination of Brazil for these practices than in Peru and the Dominican Republic, which seem to adopt them more often.

Thus, the very high or high CV classification in items D and E, which make up the self-assessment method, evinces a concern of teachers to value the participation of students, placing them as co-responsible for their training process. This shows, to a certain extent, an inverse movement to that in the previous category, which focused assessment in externally determined criteria.

Alonzo, Valencia, and Vargas (2018), for example, consider self-assessment to be a fundamental part of the holistic education of students as it promotes reflection on an individual's action in a certain activity to correct aspects in the process of their education.

Rodríguez, Ibarra, and García (2013) define self-assessment as the involvement of students in making judgments about the results of their learning, especially in formative assessments. The authors mention that self-assessment is an important requirement as it enables the development of self-perception skills and the self-regulation of one's own learning.

However, it is necessary to problematize the evaluative conception that has guided the self-assessment methods employed by teachers as PISA has emphasized socio-emotional competencies (Fernandes *et al*., 2018; Sassaki *et al*., 2018; Miranda *et al*., 2015), which may involve both individual and peer self-assessments. In this context, self-assessment could also be aligned with standardized tests such as PISA, influencing teachers' assessments.

Analysis shows that the Dominican Republic has the highest average in the feedback category, whereas Chile and Brazil, the lowest ones. Data variability suggests that teachers from the five countries indicate diverse preferences and approaches when providing feedback to students during the evaluation processes. We found a greater convergence in the responses of teachers in Panama and the Dominican Republic regarding specific feedback practices.

This category evinced evaluative methods that focus on feedback, considered essential for teachers to regulate their teaching processes and for students to continuously and dialogically become aware of their learning.

Biggs (2005) considers feedback assessment to be an action that can find what and how students learn given that, on many occasions, students, rather than studying to learn a certain content, are simply motivated to obtain a good grade, nullifying the possibility of adopting strategies to change this situation. An evaluation process that takes feedback as a basic principle enables students to be taught how they are developing learning activities, how they can do better, what are the most common mistakes they make, and how they can overcome them. The development of systematic feedback also makes it possible to provide students with metacognitive processes as students can even self-evaluate their performance from the information their teachers provided them (Biggs, 2005; Stieg, 2022).

It is interesting to note that only Brazil has a very high CV in item G, whereas countries such as Panama and the Dominican Republic have an average CV. We understand that the lack of a feedback immediately evinces a concept of evaluation as a product, an end in itself (Stieg, 2022). Therefore, evaluation loses its main potential, which is to guide teachers' pedagogical practice to guarantee the right to learning.

Given this scenario, PISA results should help teachers to rethink the evaluation in classrooms. According to Macias and Monereo (2018), this evaluation should be an instrument of constant feedback to educational work, mainly exercising the substantial function of evaluation, rather than only as a measurement instrument.

For Vilca Cruz *et al*. (2022), Peruvian schools seem to have a minority number of teachers who have effectively employed feedback to guide students who can learn and generate meaningful and good-quality knowledge so they can perform their social role in an integral and autonomous way in and for life. The authors add that feedback is necessary to overcome what is not learned or what remains to be achieved by the adequate perception of teachers, who must have the ability to administer and manage this preponderant factor for the proper functioning of formative assessment with positive results (Vilca Cruz *et al*., 2022).

Montalván (2017) highlights the positive effect feedback spaces have inside and outside classrooms. For the author, these actions enable students to feel more confident, making the climate

conducive to learning and their development toward educational objectives. On this point, we agree with Montalván (2017): teachers' help and timely action in the teaching-learning process enable students to achieve the skills established in the objectives of the educational processes.

However, it is necessary to mention what is expressed by the teachers' answers to the PISA 2018 contextual questionnaire — which give more interest to promoting evaluation methodologies focused on hetero-assessment —, but it is also necessary to enhance the use of self-assessment and feedback methodologies.

The analysis of the three categories (hetero-assessment, self-assessment, and feedback) shows a disparity in teachers' responses in the five countries in this study. However, despite these three evaluation methods, items that support the ideas of peer evaluation, known as co-assessment (Castejón Oliva *et al.*, 2011), and shared evaluation, which enables dialogue and decision-making in evaluation processes remain lacking (López-Pastor, 2009).

In fact, the inclusion of items in the PISA contextual questionnaire that also emphasize these two approaches to the development of learning assessment methods in basic education (co-assessment and shared assessment) can significantly contribute to the collection of specific data on the mutual assessment relationships established between students, explaining the meanings of how students apply what they learn (Santos, 2005). It is important to highlight that these results may reflect contextual differences, especially those related to political, cultural, social aspects, and geographical dimensions.

Conclusions

This study, which aimed to analyze how often teachers in five Latin American countries (Brazil, Chile, Panama, Peru, and the Dominican Republic) use certain methods of assessing their students' learning, indicated some similarities and distances. Thus, the methods of learning assessment used by teachers from different Latin American countries include hetero-assessment, self-assessment, and feedback.

The results in the first category showed that the practice of developing and applying own assessments varied considerably between countries. Panama stands out for its relatively lower variability, indicating greater consistency in responses. The application of standardized tests also reflects differences in preferences, with Brazil and Panama showing less variability. Chile and the Dominican Republic have a more marked preference for this practice. The methods of collecting data from class or homework assignments show variations, with Panama again showing less variability in responses.

According to teachers' narratives, the second category evinced that self-assessment methodologies are more diverse in Brazil than in other countries. While the practice of having students answer questions before their class shows a positive overall trend, the variability in responses highlights significant differences. Students' self-assessment also reflects marked differences between countries, with Brazil showing the greatest variability, indicating a diversity of perspectives among teachers. This heterogeneity suggests the need for a deeper reflection on the

clarity of the criteria of the evaluation and the understanding of their uses by teachers, stressing the importance of considering not averages and the dispersion of answers for a more comprehensive understanding of these evaluation practices in the educational context.

In the third category, the practice of providing written feedback to students' work and a grade shows variability. Brazil showed the greatest. The observation of students during particular tasks and feedback also varies, with Panama and the Dominican Republic showing less variability.

General trends show that Brazilian teachers evinced greater variability in their answers, indicating a diversity of methodologies to assess students' learning, including from those with characteristics of hetero-assessment, self-assessment, and feedback. Chile and the Dominican Republic tend to prefer standardized tests, but with variations in preference for other practices. Panama indicated greater consistency in responses, indicating a more uniform approach to evaluation practices, whereas Peru showed variation in preferences, tending toward self-assessment and standardized tests (hetero-assessment).

Analysis of these three categories show that the evaluation methods indicated by teachers vary substantially between countries. Hetero-assessment remains predominant, reflecting a traditional approach centered on teachers' active role. Self-assessment and feedback, although present, show significant variability, indicating different degrees of adoption by teachers.

The absence of items that support peer review (co-assessment) and shared assessment highlights a gap in the assessment practices considered in this study. The inclusion of these elements in the PISA 2018 contextual questionnaire can enrich the understanding of the evaluation relations established between students and teachers, viewing learning processes more comprehensively and collaboratively.

Results also highlight the importance of considering contextual differences, including political, cultural, social, and geographical aspects, when interpreting evaluation practices. Each country has particularities that influence teachers' choices in the development and application of assessment methods. Understanding these results is important to assist educational policies and pedagogical practices that are more in line with local needs and preferences. It also stresses the importance of contextualization when interpreting educational data since the methodologies indicated in the PISA contextual questionnaire showed significant variations across contexts.

When questioning about the possible impacts of PISA on teachers' evaluation practices, we observed the articulation between the discourses of international organizations and the evaluation methods indicated by teachers, which highlights the complexity of the Latin American educational scenario. While the search for global standards is evident, the practical implementation of these standards varies substantially between countries.

The evaluation methods of the teachers who responded to the 2018 PISA contextual questionnaire showed similarities. Comparative analysis indicated the complexity and diversity of evaluation practices in education, highlighting the need for a more comprehensive approach that encompasses different evaluation modalities and fosters collaboration among students. Understanding these aspects can contribute to improving educational practices and promoting more effective and inclusive learning environments.

Importantly, these findings are based on the data provided and preliminary analyses of the 2018 PISA contextual questionnaire, but a more detailed investigation is needed for a full understanding of the education landscape in each country. This comparative analysis overviewed the trends in the items in five countries, but a more in-depth analysis can be carried out with other statistical methods and from other categories.

References

ALONZO, Diana Lizbeth Rivera; VALENCIA, Marvel del Carmen Gutiérrez; VARGAS, Jorge Albino Contreras. La habilidad de autoevaluación de hábitos de estudio. *Revista Varela, Santa Clara*, v. 18, n. 49, p. 69-81, 2018. https://revistavarela.uclv.edu.cu/index.php/rv/article/view/81

BARTHOLOMEU, Daniel; MONTIEL, José Maria; NÉIA, Soraia; SILVA, Marjorie Rocha. Habilidades sociais e desempenho escolar em português e matemática em estudantes do ensino fundamental. *Temas em Psicologia*, v. 24, n. 4, p. 1343-1358, 2016. https://doi.org/10.9788/TP2016.4-09Pt

BIGGS, John. *Calidad del aprendizaje universitario*. Madrid: Narcea, 2005.

BLACK, Paul; WILIAM, Dylan. Assessment for learning in the classroom. *In:* GARDNER, John. *Assessment and learning*. London: SAGE Publications, 2006. p. 9-25.

BLOCH, Marc. Para uma história comparada das sociedades europeias. *In:* BLOCH, Marc *História e historiadores*. Lisboa: Teorema, 1998. p. 119-150.

BUSCÀ, Francesc; PINTO, Patrícia; MARTÍNEZ, Lurdes; PEIRE, Tomás. Sistemas y procedimientos de evaluación formativa en docencia universitaria: resultados de 34 casos aplicados durante el curso académico 2007-2008. *Estudios sobre Educación*, n. 18, p. 255-276, 2010. https://doi.org/10.15581/004.18.4674

CARDONA, Juan Ricardo. *Evaluación educativa*: una herramienta para la toma de decisiones. Madrid: Magisterio, 1994.

CARNOY, Martin; KHAVENSON, Tatiana; FONSECA, Izabel; COSTA, Leandro; MAROTTA, Luana. A educação brasileira está melhorando? Evidências do PISA e SAEB. *Cadernos de Pesquisa*, v. 45, n. 157, p. 450-485, 2015. https://doi.org/10.1590/198053143331

CARVALHO, Luís Miguel. PISA, política e conhecimento em educação. *Educação & Sociedade,* v. 37, n. 136, p. 601-607, 2016. https://doi.org/10.1590/ES0101-73302016168897

CASANOVA, María Antonia. *Manual de evaluación educativa:* escuela básica. Madrid: Muralla, 1997.

CASTEJÓN OLIVA, Francisco Javier; LÓPEZ-PASTOR, Víctor Manuel; JULIÁN CLEMENTE, José Antonio; ZARAGOZA CASTERAD, Javier. Evaluación formativa y rendimiento académico en la formación inicial del profesorado de educación física*. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, v. 11, n. 42, p. 328-346, 2011. https://www.redalyc.org/pdf/542/54222171007.pdf

CASTILLO, Santiago Arredondo; CABRERIZO, Jesús Diago. *Evaluación educativa de aprendizajes y competencias*. Madrid: Pearson, 2012.

CONTRERAS, Johana; TORRES, Álex. Las evaluaciones educacionales estandarizadas desde la experiencia de los actores. *Educação & Pesquisa*, v. 49, p. 1-19, 2023. https://doi.org/10.1590/S1678-4634202349248451es

COTTA, Tereza Cristina. Avaliação educacional e políticas públicas: a experiência do Sistema Nacional de Avaliação da Educação Básica (SAEB). *Revista do Serviço Público*, v. 52, n. 4, p. 89-111, 2001. https://doi.org/10.21874/rsp.v52i4.316

CRESWELL, John; PLANO CLARK, Vicki. *Designing and conducting mixed methods research*. Thousand Oaks: Sage Publications, 2013.

ESTEBAN, Maria Teresa. Exigências democráticas/exigências pedagógicas: avaliação. *Tecnologia Educacional*, v. 29, n. 148, p. 3-6, 2000.

FERNANDES, Luana de Mendonça; LEME, Vanessa Barbosa Romera; ELIAS, Luciana Carla dos Santos; SOARES, Adriana Benevides. Preditores do desempenho escolar ao final do ensino fundamental: histórico de reprovação, habilidades sociais e apoio social. *Trends in Psychology*, v. 26, n. 1, p. 215-228, 2018. https://doi.org/10.9788/TP2018.1-09Pt

GINZBURG, Carlo. *Mitos, emblemas e sinais:* morfologia e história. São Paulo: Companhia das Letras, 1989.

GOMES, Frederico Pimentel. *Curso de estatística experimental*. São Paulo: Nobel, 1985.

GOSÁLBEZ, Alfredo Celdran. *Técnicas para el tratamiento psicopedagógico*. Madrid: Kapelusz, 1989.

JOLANDEK, Emilly Gonzales; PEREIRA, Ana Lúcia; MENDES, Luiz Otavio Rodrigues. Avaliação em larga escala e currículo: relações entre o PISA e a BNCC. *Com a Palavra, o Professor*, v. 4, n. 10, p. 245-268, 2019. https://doi.org/10.23864/cpp.v4i10.370

LÓPEZ-PASTOR, Víctor Manuel. *Evaluación formativa y compartida en educación superior:* propuestas, técnicas, instrumentos y experiencias. Madrid: Narcea, 2009.

MACIAS, Esperanza Mejías; MONEREO, Carlos Font. Transformar a avaliação através do PISA: o projeto GAPPISA. *In:* ORTIGÃO, Maria Isabel Ramalho. *Políticas de avaliação, currículo e qualidade:* diálogos sobre o PISA. Curitiba: CRV, 2018. p. 69-90.

MANN, Henry; WHITNEY, Donald. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics,* v. 18, n. 1, p. 50-60, 1947. https://doi.org/10.1214/aoms/1177730491

MATOS, Daniel Abud Seabra; FERRÃO, Maria Eugénia. Repetência e indisciplina: evidências de Brasil e Portugal no Pisa 2012. *Cadernos de Pesquisa*, v. 46, n. 161, p. 614-636, 2016. https://doi.org/10.1590/198053143669

MIRANDA, Gilberto José; LEMOS, Karine Custódio da Silva; OLIVEIRA, Allana Santos; FERREIRA, Mônica Aparecida. Determinantes do desempenho acadêmico na área de negócios. *Revista Meta: Avaliação*, v. 7, n. 20, p. 175-209, 2015. https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/264

MONTALVÁN, Denisse Albuquerque. *Características asociadas a la evaluación formativa y su relación con el aprendizaje de habilidades matemáticas en estudiantes de la carrera profesional de administración de empresas de la Universidad Privada SISE, 2016*. Dissertação (Máster en Ciencias de la Educación con Mención en Educación Matemática) – Universidad Nacional de Educación Enrique Guzmán y Valle, Lima, Peru, 2017.

OZGA, Jenny. The politics of accountability. *Journal of Educational Change*, v. 21, n. 1, p. 19-35, 2020. https://doi.org/10.1007/s10833-019-09354-2

PINTO, Joaquim; BIXIRÃO NETO, Teresa; CARVALHO, Jaime e Silva. Fatores influenciadores do desempenho de estudantes portugueses, singapurenses, holandeses, espanhóis e brasileiros em literacia matemática no

PISA: revisão integrativa. *REXE: Revista de Estudios y Experiencias en Educación,* v. 18, n. 37, p. 41-60, 2019. https://doi.org/10.21703/rexe.20191837dapiedade7

RAMOS-ZINCKE, Claudio. Dispositivo de evaluación y gubernamentalidad del sistema educacional: entretejimiento de ciencia social y poder. *Cinta de Moebio,* Santiago de Chile, v. 61, p. 41-55, 2018. http://dx.doi.org/10.4067/S0717-554X2018000100041

RODRÍGUEZ, Gregorio Gómez; IBARRA, María Soledad Sáiz; GARCÍA, Eduardo Jiménez. Autoevaluación, evaluación entre iguales y coevaluación: conceptualización y práctica en las universidades españolas. *Revista de Investigación en Educación*, v. 11, n. 2, p. 198-210, 2013. http://webs.uvigo.es/reined/

SALOKANGAS, Maija; KAUKO, Jaakko. Tomar de empréstimo o sucesso finlandês no PISA? Algumas reflexões críticas da perspectiva de quem faz este empréstimo. *Educação e Pesquisa,* v. 41, p. 1353-1366, 2015. https://doi.org/10.1590/S1517-9702201508144214

SANTOS, Wagner dos. *Currículo e avaliação na educação física*: do mergulho à intervenção. Vitória: Proteoria, 2005.

SASSAKI, Alex Hayato; DI PIETRA, Giovanni; MENEZES FILHO, Naercio; KOMATSU, Bruno. Por que o Brasil vai mal no PISA? Uma análise dos determinantes do desempenho no exame. *Policy Paper*, n. 31, p. 1-24, 2018.

SCHLEICHER, Andreas. *PISA 2018:* Insights and interpretations. Paris: OECD Publishing, 2019.

SOARES, José Francisco; CANDIAN, Juliana Frizzoni. O efeito da escola básica brasileira: as evidências do PISA e do SAEB. *Revista Contemporânea de Educação*, v. 2, n. 4, p. 163-181, 2007. https://revistas.ufrj.br/index.php/rce/article/view/1522

STIEG, Ronildo. *Autores, teorías y concepciones de evaluación*: un análisis comparado en la formación de profesores de educación física en siete países de América Latina hispanohablantes. Tese (Doutorado em Educação Física) – Programa de Pós-Graduação em Educação Física, Universidade Federal do Espírito Santo, Vitória, 2022.

STIEG, Ronildo; FERREIRA NETO, Amarílio; FROSSARD, Matheus Lima; SANTOS, Wagner. A prescrição para o ensino das tipologias de avaliação na formação de professores em educação física na América Latina. *Currículo sem Fronteiras*, v. 22, p. 1-33, 2022. http://dx.doi.org/10.35786/1645-1384.v22.1872

TAMAYO, Rafael Lodezma Caballero. Heteroevaluación, autoevaluación, coevaluación y evaluación compartida: consideraciones conceptuales. *Revista Atlante Cuadernos de Educación y Desarrollo*, p. 1-16, 2018. https://www.eumed.net/rev/atlante/2018/10/heteroevaluacion-autoevaluacion.html

VANLOMMEL, Kristin; SCHILDKAMP, Kim. How do teachers make sense of data in the context of high stakes decision making? *American Educational Research Journal*, v. 56, n. 3, p. 792-821, 2019. https://doi.org/10.3102/0002831218803891

VERA, José Ángel Noriega; BUENO, Gissel Castro; CALDERÓN, Nohemí Guadalupe González; MEDINA, Francisca Leticia Figueroa. Modelo de autoevaluación y heteroevaluación de la práctica docente en Escuelas Normales. *Educação & Pesquisa*, , v. 44, p. 1-19, 2018. https://doi.org/10.1590/S1678-4634201844170360

VERGER, Antoni; FONTDEVILA, Clara; PARCERISA, Luís. Reforming governance through policy instruments: how and to what extent standards, tests and accountability in education spread worldwide. *Discourse: Studies in the Cultural Politics of Education*, v. 40, n. 2, p. 248-270, 2019. https://doi.org/10.1080/01596306.2019.1569882

VILCA CRUZ, Abigail Yessica; HUARANCCA RAMÍREZ, Delfina; MAMANI CRUZ, Ivonne; APAZA RIVAS, Edwin; CONTRERAS RIVERA, Robert Julio. La retroalimentación formativa, un factor clave del aprendizaje

matemático en la educación básica primaria. *Ciencia Latina Revista Multidisciplinar,* v. 6, n. 6, p. 7274-7288, 2022. https://doi.org/10.37811/cl_rcm.v6i6.3945

_____

RONILDO STIEG

PhD in Physical Education, Universidade Federal do Espírito Santo (UFES), Vitória, Espírito Santo, Brazil; Postdoctoral Fellow in Education, Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES), Vitória, Espírito Santo, Brazil.

DENILSON JUNIO MARQUES SOARES

PhD in Education, Universidade Federal do Espírito Santo (UFES), Vitória, Espírito Santo, Brazil; Professor, Instituto Federal de Minas Gerais (IFMG), Ouro Preto, Minas Gerais, Brazil.

WAGNER DOS SANTOS

PhD in Education, Universidade Federal do Espírito Santo (UFES), Vitória, Espírito Santo, Brazil; Director of Postgraduate Studies, Universidade Federal do Espírito Santo (UFES), Vitória, Espírito Santo, Brazil.

AUTHORS' CONTRIBUTION

Author 1 – Analysis and interpretation of data; preparation of the final text.
Author 2 – Construction and processing of data; preparation of the final text.
Author 3 – Conception and design of the research; preparation of the final text.

AVAILABILITY OF RESEARCH DATA

Data will be provided upon request.

HOW CITY THIS ARTICLE

STIEG, Ronildo; SOARES, Denilson Junio Marques; SANTOS, Wagner dos. Assessment methods for learning: what do teachers from Latin American countries show in the PISA 2018 contextual questionnaire? *Educar em Revista,* Curitiba, v. 40, e93764, 2024. https://doi.org/10.1590/1984-0411.93764

This article was reviewed by Tikinet Edição Ltda. – EPP – E-mail: comercial@tikinet.com.br. After being designed, it was submitted for validation by the author(s) before publication.

_____