

Literacy assessment in Espírito Santo: analysis of Paebes Alfa between 2009 and 2017¹

Avaliação da alfabetização no Espírito Santo: uma análise do Paebes Alfa entre 2009 e 2017

Wagner Silveira Rezende*

ABSTRACT

This article analyzes Paebes Alfa, the literacy assessment program of the state of Espírito Santo, between 2009 and 2017. Therefore, a set of ten criteria was proposed for analysis and description of the program, something not found in the educational literature, whose mapping showed the absence of studies about this theme. The study included detailed bibliographic and documentary research, as well as interviews with managers and specialists who worked in the program throughout the analyzed period. It is expected that this work will be able to serve as a starting point for researches about literacy assessment programs in Brazil, helping to fill the current gap in the specialized literature.

Keywords: Literacy assessment. Paebes Alfa. Educational assessment.

RESUMO

Este artigo faz uma análise do Paebes Alfa, o programa de avaliação da alfabetização do estado do Espírito Santo, entre os anos de 2009 e 2017. Para tanto, foi proposto um conjunto de dez critérios para análise e descrição do programa, algo não encontrado na literatura educacional, cujo mapeamento mostrou a ausência de estudos dessa natureza. O estudo contou com detalhada pesquisa bibliográfica e documental, além de entrevistas com gestores e

¹ Translated by Michelle de Abreu Aio. E-mail: aiomichelle@gmail.com

* Universidade Federal de Juiz de Fora. Juiz de Fora, Minas Gerais, Brasil. Email: wagner.rezende@ufjf.edu.br - <https://orcid.org/0000-0003-1617-282X>

especialistas que trabalharam com o programa ao longo do período analisado. Espera-se que este trabalho seja capaz de servir como ponto de partida para pesquisas que se interessam pela análise e compreensão de programas de avaliação da alfabetização no Brasil, ajudando a preencher a atual lacuna observada na literatura especializada.

Palavras-chave: Avaliação da alfabetização. Paebes Alfa. Avaliação educacional.

Introduction

This article is derived from the doctoral research entitled “*Avaliação em larga escala da alfabetização: os casos de Paebes Alfa, Proalfa² e Spaece Alfa³*” [Large-scale assessment of literacy: The cases of Paebes Alfa, Proalfa and Spaece Alfa] (REZENDE, 2020), defended in May 2020⁴. Its aim was to analyze these three state literacy assessment programs in detail in order to contribute to filling the gap found in the educational literature regarding assessment policies for this stage of schooling. The importance of educational assessment systems (CASTRO, 2009) and the central role of policies related to literacy (GONTIJO, 2014) contrast with the lack of interest in educational literature on literacy assessment policies embodied in state programs like the three analyzed in the thesis. Such contrast becomes even more evident when we observe the importance that state managers have attributed to these programs, as will be shown later. In the available literature, no studies were found that made a comprehensive analysis of programs such as Paebes Alfa (*Programa de Avaliação da Educação Básica do Espírito Santo – Alfabetização*)⁵, which performs a systematic assessment of literacy in the state of Espírito Santo (ES).

Part of the literature on public policies (FREY, 2000; MARTINS, 2013) criticizes the excess of papers describing specific programs, drawing attention to the need to establish links with the macro-political context, especially the international one. However, regarding literacy assessment programs, we observe

2 Proalfa, in Minas Gerais.

3 Spaece Alfa, in Ceará.

4 Education Graduation Program of *Universidade Federal de Juiz de Fora* under advisory of Professor Hilda Micarello.

5 Conducted by the State Education Department.

the opposite – the lack of studies that describe and analyze these programs in depth. The few papers that address Paebes Alfa, for example, generally do so by taking the program as a source of data, but not as an object of study, or from a critical bias (affecting, for example, the pedagogical practice in schools) without carrying out a deeper analysis and description.

This article presents the results of the research on Paebes Alfa between 2009 and 2017 (covering almost its entire history). To do so, a search was carried out in three important national journals and in *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (Capes) Theses and Dissertations Database (presented in the following topic) in order to investigate how the program has been portrayed in the educational literature. Then, the criteria proposed for the analysis of the program are presented. They were specifically designed for the analysis of literacy assessment policies – something that has not been found in the literature. Thus, the research included a detailed analysis of documents related to the program (result disclosing materials, Education Department's documents, contracts, records of the partner institution for implementing the program etc.) and interviews with two managers who were ahead of Paebes Alfa during the period selected for analysis, as well as experts from the *Centro de Políticas Públicas e Avaliação da Educação* [Public Policies and Education Assessment Center] da *Universidade Federal de Juiz de Fora* (CAEd/UFJF) who participated in the program implementation (REZENDE, 2020).

The Paebes Alpha in Educational Literature

The research in the national literature sought to confirm the hypothesis that literacy assessment programs did not arouse the interest of research in education in the country, although they have been given some importance in the scenario of state educational policies. To this end, three important journals in the field of education were consulted between 2004 and 2017: a) *Estudos em Avaliação Educacional* (EAE), by *Fundação Carlos Chagas* (FCC), a journal specialized in topics related to educational assessment (one of the most important in the country in this area – Qualis A2); b) *Revista Brasileira de Educação* (RBE), selected because it is an *Associação Nacional de Pós-Graduação e Pesquisa em Educação* (ANPEd) journal, an important vehicle in Brazil for disseminating educational research on a wide range of topics (Qualis A1); and c) *Revista Brasileira de Estudos Pedagógicos* (RBPE), selected for being published by *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira* (INEP),

the body responsible for conducting national assessments (Qualis B1). For the articles to be considered, the research was carried out as follows: The words *assessment* and/or *literacy* should be in the paper's title or keywords (or in the names of assessment programs, such as *Provinha Brasil*, *Avaliação Nacional da Alfabetização* (ANA), *Spaace Alfa*, *Paebes Alfa*, *Proalfa*, etc.), or the content in the abstract should be directly linked to the subject of literacy assessment.

In EAE, 12 of the 404 articles were dedicated to the assessment of literacy (2.9% of the total). Of these, *Paebes Alfa* is mentioned in only one of them (ROCHA; MARTINS, 2012). The authors' paper proposes a children's writing corpus using data from *Paebes Alfa* from the 2010 and 2011 editions. A brief reference to the program is made in the abstract, but it is just the source of the data, not the object of study by the authors. In RBE, four of the 668 papers were on literacy assessment (0.59% of the total). None of them dealt with *Paebes Alfa*, and three were related to *Provinha Brasil*: (ESTEBAN, 2012; MORAIS, 2012; LEAL *et al.*, 2017). In RBPE, only one of the 473 papers dealt with the topic of literacy assessment (0.21% of the total), once again with a focus on *Provinha Brasil* (FREITAG; ALMEIDA; ROSÁRIO, 2013).

Thus, the articles dealing with literacy assessment in the researched periodicals were mainly about *Provinha Brasil*. There were also articles on *Sistema Permanente de Avaliação da Educação Básica do Ceará – Alfabetização* (*Spaace Alfa*), *Programa Alfabetização na Idade Certa* (*Paic*) and its assessment model in Ceará, and on the construction of a proficiency scale for assessing youth and adult literacy. Those three central journals for the dissemination of educational research in Brazil present a framework that highlights the limited production on literacy assessment in the country in a period that involves the emergence and development of the main assessment programs for this stage in the states. On the three journals, 17 papers were published on the subject (only one mentions *Paebes Alfa*, but without taking it as a research object) among 1,545 articles, which corresponds to 1.1% of the total.

We might think that little has been published in these journals, though it does not necessarily mean that little was written about the topic. Thus, the decision was made to also analyze Capes' Theses and Dissertations Database (consultation in July 2019) to verify whether *Paebes Alfa* was an object of interest for graduate studies in the country. Due to the characteristics of the Capes database search tool, the search for *Paebes Alfa*, in these terms, was not possible as the word "*Alfa*" led to a huge number of papers that used the term *alfabetização* [literacy] in their titles or keywords. Therefore, the search was made with the term *Paebes*. The results from this search would bring papers both related to *Paebes* and *Paebes Alfa* and the selection of the latter would be made on a case-by-case basis.

There were ten papers as a result of this search, four of them directly or indirectly related to Paebes Alfa. This is a timid result for a program that has been around for over a decade. Of the four papers, two of them deal specifically with Paebes Alfa (CÔCO, 2014; SILVA, 2013), one of them compares Paebes Alfa with other programs, while the other studies a specific topic based on data from Paebes Alfa, without the program itself being its main object of study. The latter is the thesis by Lima (2016), which has educational inequalities, more specifically the inequality of knowledge among students, as its central topic. The author uses data from Paebes Alfa to propose a debate on educational inequalities in the national scenario.

Côco's (2014) thesis has Paebes Alfa as its object of study, analyzing the program regarding the production of meanings for this evaluation and its relationship with literacy practices in 1st- and 2nd-year classes in the municipal system of Serra, in Espírito Santo. The way it generates effects on literacy practices in the classroom is an important feature. However, it does not address the program, but some of its characteristics. Her interest is in the classroom effects and not in the program itself. Silva's (2013) thesis analyzes the results of writing at the end of the 2nd year of elementary school using the Paebes Alfa as a reference. The paper is dedicated to a technical analysis of the reference matrix and the correction keys of the writing assessment, culminating with the proposition of a new correction key for the items. Once again, it is not the program that is the object of research, but the data it provides.

Portela's (2016) dissertation is dedicated to analyzing the reference matrices of three literacy assessment systems, including the Paebes Alfa – the other two are *Sistema de Avaliação Educacional de Pernambuco* (Saepé) and *Sistema Estadual de Avaliação da Aprendizagem Escolar* (Seape). The author's aim is to compare the reference matrices of the three programs with ANA's matrix and propose a unified matrix for mathematics assessments in the 3rd year of elementary school. It is a paper that analyzes an important dimension of Paebes Alfa, but it is not a study about the program itself.

From the other six papers that came from the search, four of them involve Paebes (PEREIRA, 2014; OLIVEIRA, 2018; PAULA, 2018; OLIVEIRA JÚNIOR, 2013) and do not mention Paebes Alfa (the other two do not even mention Paebes or Paebes Alfa). Therefore, by observing the educational literature consulted we understand that Paebes Alfa is practically invisible as a research object.

The Selected Criteria for the Analysis of Paebes Alfa

The analysis of public policies often uses typologies definition. Based on the criteria selection, policies are categorized into different types, allowing for investigations of different natures (comparisons between types, policy effectiveness, side effects, adjustment to certain types of political context, etc.). However, the types proposed by research on public policy can be very generic depending on the policies to be analyzed. This makes the typology innocuous because very different policies can be categorized as being of the same type, setting aside fundamental differences, which happens with the classic typology by Lowi (MAIA; GRANDO, 2016) and that by Deubel (2009). Analyzing the selected literacy assessment programs required the selection of specific criteria not found in the literature and that would allow for a more detailed understanding of each one of them, making room for a later comparative analysis. Therefore, the path chosen was the selection of 10 criteria for the analysis of literacy assessment policies embodied in state programs. The justifications and details of each criterion are found in Rezende (2020), and the table below summarizes the description of what will be analyzed in each one of them.

It is important to emphasize that this is not an exhaustive list of criteria that should be adopted to carry out a survey of this nature – i.e., interested in analyzing a state literacy assessment program in greater detail. The reason is that such criteria are important and provide a very broad view of a literacy assessment program, but other criteria could be used. As we have seen, the educational literature is lacking in research of this nature. Therefore, the criteria proposition presented here can serve as a starting point for reformulations and adjustments for research that are interested in analyzing and understanding literacy assessment programs. Following the table, Paebes Alfa is analyzed based on each of the criteria.

TABLE 1 – CRITERIA FOR A TYPOLOGY OF LITERACY ASSESSMENT POLICIES

Criteria	Basis for description
1. Scope of policy	What is the scope of the policy? Its range. Does it involve the entire educational system or is it just a part of it?
2. Target audience	Who is the policy's target audience? Who is it primarily aimed at? What are the secondary recipients?
3. Creation context (agenda) and implementation model	How was the agenda around the policy? Is it possible to track it? Is there a record of that moment? Which agents acted decisively so that the agenda could be defined? How was the policy implemented? Was the model <i>top/down</i> used? Was it <i>bottom-up</i> ? Was there street-level bureaucracy or just higher-level decision-makers involved?
4. Assessment design	What are the instruments mobilized by the policy? What grades and subjects are evaluated? What is the proficiency scale? How is it applied? What are the reference matrices?
5. Dissemination and use of results (nature of dissemination and communication channels)	Are policy results disclosed? If so, by what means and for whom? How is access to these results? What kind of information is made available? How long after the assessment are the results available? Are the delivery, understanding, and use of results dimensions considered in the disclosure?
6. Conception of literacy	What is the literacy concept that underlies the development of reference matrices and test creation?
7. Link with other educational policies and policy generation (<i>accountability</i>)	Does the policy relate to other educational policies? If so, which are they? Is there an accountability policy linked to the evaluation results? If so, of what type? Are the consequences of this policy mild or severe for its participants? To which generation of educational assessment does the policy belong?
8. Longevity (duration and constancy)	What is the constancy and duration of the policy? What transformations has it undergone over time?
9. Participation of municipalities	How is the participation of municipalities in the program? Is it done by adherence? Are expenses for participating in the program covered by the state or does the municipality also have expenses?
10. Volume and transparency of resources applied in evaluation	Are resources disclosed transparently? Is there accountability to the general population? Where is the disclosure made? What is the volume of resources allocated to the evaluation when compared to the general expenses of the states with education?

SOURCE: Author's elaboration.

Criterion 1: Paebes Alfa, since its first edition in 2008, has evaluated schools in state and municipal education systems. From the 2nd wave evaluation in 2010 (carried out at the end of that year), the program also started to evaluate private schools that wished to participate. In the scope of public systems,

assessment is carried out by census. As the program assesses the three grades of the literacy cycle, this means that in principle all students in the cycle are assessed. In the case of municipal systems, their participation occurs through adherence. Thus, the schools in the system are not required to participate in the assessment, but by participating they are assessed in the three grades. Since its second edition (2009), Paebes Alfa has had a high number of participating municipalities (76), reaching a total (78) adherence of municipalities in 2016. Regarding the private system, the decision to participate is up to the school. As the program provides for the evaluation of all students enrolled in all grades of the literacy cycle in public schools in addition to allowing the participation of private schools, we can say that Paebes Alfa has a wide reach (covers all students in the literacy cycle).

Criterion 2: When it comes to the target audience, the results of the assessment of Paebes Alfa are delivered mainly to decision-makers within the systems and schools. Thus, the education secretary and their staff in the Department, regional superintendents, principals, and teachers are the policy's immediate audience, for whom the results are designed and who are expected to use them. This can be seen in the results disclosure materials. However, although these are the target audience of the program, we can consider that society, including the students' parents, the community around the school and researchers, besides those interested in the topic, are also a target audience – after all, at the school level the results are available to any user of the program's website. Students and their guardians have already been the target audience of the evaluation results in some editions of the program, but they no longer receive specific materials. Like any other user, they can access the school results through its website, and it is expected that students results will serve as a basis for teacher planning in the classroom, especially through the disclosure of results based on *Teoria Clássica dos Testes* [Classical Test Theory] (TCT) [for more on TCT, see Pasquali (2013)]. This is the effort of the disclosure process.

Criterion 3: implementation has been treated as a key element in the assessment of public policies. Januzzi (2019), from a systemic perspective, proposes the redesign of the classic cycle of policies. To do so, implementation figures as an almost final stage of the process, relocating this stage to the center of the cycle with more realistic objectives. According to the author, it is in such stage that “programs ‘live’ most of the time and where corrections, improvements or incremental innovations will determine the success or failure of the public intervention” (JANUZZI, 2019, p. 73, our translation). Draibe (2001) goes in the same direction, emphasizing the importance of evaluating the implementation, especially in process evaluations. For the author, the strategies that involve the

implementation cannot be forgotten, as they are fundamental for “the correct identification of the process factors that operate either as facilitators or as obstacles to the program” (DRAIBE, 2001, p. 29, our translation).

In any case, it is difficult to track all the forces and influences used to create a public policy. Despite this, it is possible to identify elements that were decisive for the implementation of the program in the context of Paebes Alfa’s creation. According to Segatto (2011), the main one was the administrative reform of the state which started in 2003. The Education Department was one of the first to carry out a major reform in the state based on a result-oriented culture (SEGATTO, 2011, p. 72). At the time, choosing a secretary of education with a more technical profile contributed to the development of a result-driven culture in the department, leading to the creation of an evaluation system (Paebes and its application in literacy, Paebes Alfa). Later, it led to a high-impact *accountability* policy [Brooke (2016) presents a historical analysis of salary incentives policy in education in the state of Espírito Santo], based on salary incentives, under the responsibility of the newly created Planning and Management Council (the creation of the council itself was an important step towards the creation of the evaluation system).

Segatto (2011) also emphasizes that the reform in the education management in Espírito Santo was inspired by examples from other Brazilian states, such as São Paulo, Pernambuco, Ceará, and Minas Gerais. The last three already have their own assessment systems covering the cycle of literacy, which would also have influenced the creation of a literacy assessment program in the state. Côco (2014), on the other hand, points to the Term of Commitment of the *Programa Todos pela Educação* [All for Education Program] as a crucial factor for the creation of Paebes Alfa, which formalized the commitment of the states to achieve the goals agreed upon and established by Decree No. 6094/2007 (BRASIL, 2007). One of the goals of the term established the literacy of children up to eight years old, providing for the measurement of results by specific periodic examination. Before that, the 2025 Espírito Santo Development Plan had been published in 2006 establishing the transparency of the annual results of each school, among other aspects. It is in this context of a) administrative reform, b) the influence of examples from other states, and c) the *Todos pela Educação* Commitment that Paebes Alfa was created. Its creation is linked to a response to the demands of this new way of managing education that is strongly anchored in a result-oriented culture.

Regarding the policy implementation model, Paebes Alfa is closer to the *top/down* than to the *bottom-up* model. However, it is necessary to relativize what this expression (*top/down*) means. In a literal meaning, “from top to bottom” may suggest that the policy implementation took place in a purely

imposing manner. The Education Department was, in fact, the main driver of this process. As occurred in the context of implementing the *Sistema de Avaliação da Educação Básica* [Basic Education Assessment System] (SAEB), in the early 1990s, for Paebes Alfa there was no broad discussion about the program involving the entire state system. Taking Saeb as its main reference, the creation of state educational assessment systems, including the Paebes Alfa in Espírito Santo, was not submitted to the scrutiny of teachers and school administrators.

The policy's elaboration and implementation came from the Education Department, which may have been one of the factors of initial resistance from some education agents. However, even if initially there was no such massive participation in defining the policy, throughout its existence Paebes Alfa sought mechanisms to incorporate the actions of these agents in the policy development (until the 2015 edition, for example, part of the tests was made up of items prepared by the public system's teachers trained by CAED/UFJF) (REZENDE, 2020).

Criterion 4: the design of Paebes Alfa involves a 1st, 2nd and 3rd year census assessment in reading, writing, and mathematics. The latter started to be evaluated in the 2010 edition, whereas the former are evaluated since the first edition. This is a significant difference between Paebes Alfa and its counterparts in other states, such as Minas Gerais and Ceará, for example (REZENDE, 2020). The writing assessment has put Paebes Alfa in evidence in the scenario of literacy assessment in Brazil due to its pioneering nature. Between 2009 and 2015, the program conducted entrance (1st wave) and exit (2nd wave) assessments of students in all editions. The diagnostic character of the 1st wave was added to the summative character of the 2nd wave, turning the program into a remarkable source of information on student performance (thanks to the comparability of the entrance and exit tests due to Item Response Theory – IRT).

Test reference matrices are documents with a political character. This means that the decision on which aspect of the curriculum will be evaluated belongs to the Education Department. Nevertheless, like other state assessment systems, Paebes Alfa is conducted in partnership with CAEd/UFJF, a research institution from a public university that specializes in educational assessment and public policy. In an interview given to Rezende (2020), the specialists in instruments elaboration from CAEd/UFJF stated that elaborating the reference matrices is a joint work. Done in partnership with the Education Department, they decide what they intend to evaluate, while the CAEd experts organize the matrix in a tailored format to serve as the basis for item production and test assembly. The final word on the elaborated matrix and its format is always from the department. With the matrices undergoing few changes over the years and with the partnership already consolidated with CAEd, these validations have become less frequent in the latest editions of Paebes Alfa.

For reading, Paebes Alfa presented a single matrix for the entire literacy cycle (involving the three grades evaluated) up to the 2011 edition. As of 2012, there was a process for specifying the levels for each descriptor, accompanied by the definition of which of them would be part of the specific tests for each grade evaluated, and this format was maintained in the 2013 edition. From 2014 to 2017, the matrices were produced for each evaluated grade (the three matrices – 1st, 2nd, and 3rd years – represent subsets of the same set of descriptors; the differences are present in the item classes and their difficulties) and composed by five topics, remaining unchanged throughout this period: i) Recognition of the alphabetic system convention; ii) Appropriation of the alphabetic system; iii) Reading: Understanding, analysis, and evaluation; iv) Social uses of reading and writing, and v) Written production.

With regards to writing, the 2009 matrix had two descriptors: Writing words and writing sentences/texts. As of the 2010 edition, there was a change in compliance with the complexity differences of writing sentences and texts, causing the matrix to have three descriptors (writing words, sentences, and texts) which remained until 2017. With a stable writing matrix, the interview with two specialists in writing assessment from CAEd (REZENDE, 2020) revealed that the biggest change in the writing tests of Paebes Alfa occurred in elaborating the items. The change in the partnership of the department from *Centro de alfabetização, leitura e escrita* [Center for literacy, reading and writing] da *Universidade Federal de Minas Gerais* (Ceale/UFMG) to CAEd gave more autonomy and freedom to UFJF's specialists to research the writing assessment with the purpose of improving the tests (REZENDE, 2020, p. 213), which was also coupled with the creation of a CAEd's team specialized in writing assessment.

The choice of words and imagery changed. The idea was no longer to work with figures, but with real images (photographs), in addition to prioritizing the use of cleaner images, with less secondary information, which could lead to a deviation of the student's attention. There was also a change in the command of items, which became more directive and clearer, and in the correction process (the selection criteria for graders became more rigorous). The longevity and consolidation of the writing assessment turn Paebes Alfa into a reference program regarding this topic, anticipating many trends in the area.

Until the 2012 edition, the Paebes Alfa math reference matrices were created in partnership with Ceale. As of 2013 until today, the preparation of matrices, items, and tests started to be done in partnership with CAEd. This led to a reformulation of the mathematics matrices at the request of the Education Department. The main change was the reduction of descriptors aimed at solving problems with the use of basic operations. Until 2012, each operation (addition, subtraction, multiplication, and division) had a descriptor in the

matrix. These 4 descriptors were grouped into two (addition and subtraction; multiplication and division), making room for the inclusion of other skills. With a less exacerbated focus on the numbering system, other areas gained more prominence, such as quantities and measurements, for example (reading the hours was a skill that gained ground).

During the research (REZENDE, 2020), five CAEd experts were interviewed on the development of instruments for the assessment of mathematical literacy. They stated that, throughout the editions of the program, the mathematics tests became easier for students of the 3rd year of elementary school, as the matrix was the same for the entire literacy cycle. The difficulty of the items was the only thing that changed in the tests. Thus, in 2017, for the 2018 edition to take effect, proper matrices were built for each grade assessed so that the tests would be more adjusted to the population that would respond to them.

The stability of the Paebes Alfa design can also be seen in the test booklet models and in the number of items per booklet. In mathematics, the number of booklets (8) and the number of items per booklet (24) were the same throughout the analyzed period. Between 2010 and 2011 there was a small change in the number of booklets for Portuguese language (from 8 to 12, which remained until 2017; there were 25 items per booklet between 2010 and 2017; in 2009, there were 24 items per booklet). The four performance standards were also unchanged in the period, both for reading and math: Pre-Basic, Basic, Proficient, and Advanced⁶.

Criterion 5: between 2009 and 2017, Paebes Alfa's dissemination system was based on virtual and physical channels. On the program's website (PAEBES, 2020), the evaluation results are available for open consultation to the public, sorted for all grades and subjects up to the school level (results per student are restricted to each school via login and password) for the public system. The results of the participating private schools are accessed only through a login and password and sent specifically to each of the schools so that they only have access to their own results. The website also makes available (as of the 2013 edition) the disclosure collections (bulletins and journals) of Paebes Alfa's results. Those collections, produced at each edition of the program, used to be printed and delivered to schools so that they could have this material at hand. The aim is not only to disseminate the results, but to encourage their use as a support for teaching planning. That has become one of the central elements of research in evaluation – the results' appropriation (CALDERÓN, 2017).

⁶ With the cut-offs varying for each of the three grades evaluated on a scale of 0 to 1000 points; for more on this topic, see Rezende (2020).

Moreover, there are support materials in the website available to any user since 2009 (except for the 2010 edition), such as support materials for dissemination workshops, videos of results' appropriation, and even test items. From 2013 to 2017, the collections presented the same structure, with journals aimed at specific audiences (for the 2017 edition, an executive summary is also available): i) *Revista do Sistema*, aimed at Education Departments and their teams, with results aggregated by the teaching superintendence focusing on the appropriation of results for decision-making at the system's level; the program offers a journal for each type of system (state and municipal); ii) *Revista do Gestor*, aimed at the principal, with the results of their own school and focusing on the appropriation of results at that level; iii) *Revista do Professor*, aimed at literacy teachers with a view to the pedagogical use of the results (Paebes Alfa published a specific journal with Portuguese language and mathematics results for each assessed grade – 1st, 2nd and 3rd years). In addition, throughout the analyzed period the program held face-to-face workshops given by CAED specialists for the dissemination and appropriation of results with a group of managers and teachers designated annually by the Education Department.

Criterion 6: The concept of literacy that underpinned the assessments of Paebes Alfa throughout the period was “literate literacy”. The initial partnership with Ceale for the elaboration of reference matrices is one of the factors that possibly led to the adoption of this concept. It was maintained after establishing the partnership with CAEd. Magda Soares, one of Ceale's founders, is an explicit reference in some of the program's results disclosure materials. The journal *Revista do Professor* for Paebes Alfa 2017 presents literacy and literate profiles as a new indicator of that edition (ESPÍRITO SANTO, 2017). Reference matrices reveal such influence by always predicting descriptors related to the appropriation of the alphabetic system and the recognition of its conventions (identifying directions of writing, the number of syllables in a word, identifying canonical and non-canonical syllables, among others). This is one of the possible reasons for the resistance of some literacy researchers [such as Gontijo (2014)] to large-scale assessment [for more about that, see Rezende (2020)].

Criterion 7: Paebes Alfa is a program that is part of Espírito Santo's evaluation system and directly associated with Paebes (responsible for evaluating the other stages). Although it was associated with other programs from the Education Department, such as those aimed at training teachers, Paebes Alfa was not directly linked to a representative program of the state's educational policy in the period analyzed, as occurred in other states [such as Spaece Alfa and Paic, in Ceará, and *Programa de Intervenção Pedagógica* [Pedagogical Intervention Program] (PIP) and *Programa de Avaliação da Alfabetização* [Literacy Evaluation Program] (Proalfa), in Minas Gerais (REZENDE, 2020)].

With regards to generating the evaluation policy, following the categorization by Bonamino and Sousa (2012), Paebes Alfa can be understood as belonging to the third generation for its results are used as the basis for a consolidated bonus policy that exists in the state since 2010 – *Bônus Desempenho* [Performance Bonus] [created by Complementary Law No. 504/2009 (ESPÍRITO SANTO, 2009)]. There is a cash award not linked to monthly remuneration that is awarded annually to education professionals and calculated based on individual (attendance and activities in the same school unit for at least 122 days a year) and collective (*índice de desenvolvimento nas escolas* [school development index] – IDE, which considers the results of the Paebes Alfa indicators). Its aim is to be a stimulus to improve the quality of education in the state with the assessment results as one of its foundations.

Criterion 8: Paebes Alfa is part of a long-term assessment policy. Since 2008 and without interruption, the program has carried out annual evaluations. The program underwent some adjustments, for example, in the evaluation design, but without any major changes. In addition to being long-lived, it is a stable program. In an interview for the research (REZENDE, 2020), a former deputy manager of educational assessment at the department stated that one of the reasons that helped to explain the continuity of Paebes Alfa was the fact that the schools have understood the purpose for assessing literacy, making use of the results in their work routines (REZENDE, 2020, p. 237). For the former planning and management advisor of the department (REZENDE, 2020), longevity was a result of sustaining through different administrations of the department the idea that measuring the results of the pedagogical action over time is fundamental for planning.

Criterion 9: Although it is a program conducted by the state Education Department, most of the students assessed by Paebes Alfa are enrolled in municipal schools. The participation of municipalities in the program is made by adherence to the format proposed by the state. As we have seen, all municipalities in the state have participated in the program for some years now. They are encouraged to participate in the interest of the state (for monitoring the learning of students who will eventually reach state schools), but there are no formal mechanisms, as in Ceará, for example (REZENDE, 2020, p. 244-245), to compel them to join the program (the same occurs with private schools that wish to participate). The costs of the evaluation belong to the state, the municipalities bear no expense. The cooperation term does not foresee the transfer of financial resources between the entities, stating that the cooperation has a technical nature. The state is responsible for training a coordinator responsible for evaluation in the municipality, and for ensuring that all schools participate in it. The municipalities are responsible for organizing the assessment in their domains, including the

designation of a technician who is an interlocutor with the state department, and the designation of test administrators in schools.

Criterion 10: Through Espírito Santo's government transparency website (ESPÍRITO SANTO, 2019) and through the disclosure of the contracts established with CAEd, it was possible to identify the state's expenses with the educational evaluation and to compare them with the total expenses of the Education Department throughout the analyzed period [more details on this analysis can be found in Rezende (2020)]. Expenditures on assessments refer to the assessment system as a whole, including Paebes (specific expenditures with Paebes Alfa were not detailed, as the contracts comprise the assessment of both Paebes and Paebes Alfa). It is important to emphasize that the department's total expenses account for the amounts committed (budget reserved to meet the commitments made to creditors and suppliers). Likewise, the expenses with the evaluation obtained through the contracts with the CAEd refer to the forecast of the contract, without considering any deletions or amendments.

Between 2009 and 2017, the state of Espírito Santo invested BRL 43,910,151.68 in its evaluation policy, while the total expenses of the Education Department, for the same period, were of BRL 8,017,848,181.51. Considering the entire period, the expenses with the evaluation policy account for 0.55% of the department's total expenses. In 2016, the year in which this percentage was the highest, it accounted for 0.89% of the total. Close to the importance that managers give to the program, as we will see below, expenditures do not seem to be as substantial as the criticism of external evaluation tends to consider, suggesting that the money allocated to the evaluation would be better used to solve more important problems (BARBOSA; VIEIRA, 2013, p. 429-430).

The two interviewees from the Espírito Santo's Education Department (REZENDE, 2020), who occupied important positions at Paebes Alfa throughout the period analyzed (educational evaluation management and planning and strategic management assistance), stated that the program assumed a strategic role in the state, guiding important actions in the department, such as teacher training. It also strengthened the cooperation between the state and municipalities, which adhered to the program in favor of improving the quality of education, and was the main tool in a results management context, monitoring the achievement of the goals of the national and state plans. The results started to connect principals and teachers, integrating the work of schools, and providing evidence for planning, while also evidencing an improvement in the quality of education offered. Paebes Alfa results over time confirm this improvement (REZENDE, 2020).

Paebes Alfa is, therefore, a wide-ranging program, evaluating virtually all students enrolled in the state's literacy cycle (the three grades of the state and municipal systems) and still allowing the participation of the interested private schools. Its target audience is primarily teams from Education Departments, superintendencies, school principals, and teachers, but with the school data being of public access. The context of its creation was characterized by the administrative reform of the state with a focus on results, the Commitment to All for Education, and the examples of reforms in other states (such as Minas Gerais and Ceará), whose implementation was conducted by the Education Department.

It is a census assessment of the 1st, 2nd, and 3rd years of elementary school in reading, writing, and mathematics (which places the Paebes Alfa in a prominent position when compared to other literacy assessment programs), with a great stability in relation to matrices, test booklet models and performance standards. The disclosure of its results is based on the program's website, with public access and availability of results by school, in addition to a collection of printed materials designed for specific recipients (department's staff, school management, and teachers), also available on the website, and in-person workshops for the dissemination and appropriation of results.

The concept of literacy that underlies the program is literate literacy, which is strongly influenced by the ideas of Magda Soares (2004; 2007). The program is linked to Paebes, another assessment program in the state, and to *Bônus Desempenho*, a bonus policy that is based, among other criteria, on the results of Paebes Alfa, allowing us to characterize the program as belonging to the third generation of educational assessments. It is a long-lasting (since 2008) and continuous (without interruptions) program. The participation of municipalities occurs by adherence, with all costs covered by the state and favoring cooperative federalism between entities. The average expenditure on evaluation in the state, for the period analyzed, accounts for 0.55% of the total expenditure of the Education Department, which leads us to question the criticism that the expenditure on evaluation is substantial, especially when we consider the centrality attributed to it by state managers.

Final considerations

This article allows us to make considerations about Paebes Alfa, the object of the research presented here, but also to extrapolate the analysis of the program, leading to general considerations about research on literacy assessment in the

country and educational assessment in general. The criteria used for analyzing Paebes Alfa do not exhaust all the characteristics of a literacy assessment program. The list could be different, with other characteristics being observed, or even analysing those characteristics (for example, in relation to its resources, the investigation could be on their source, whether or not there is partnership with international bodies, etc.).

Despite this, even if in a limited way, it is hoped that the analysis presented here will be able to stimulate other researchers interested in a more thorough understanding of literacy assessment programs in Brazil. As we have seen, Paebes Alfa has been analyzed by few academic papers. When analyses took place, they focused on specific aspects of the program without a more general view of it. This paper does not completely fill the gap in studies on literacy assessment programs in Brazil but seeks to shed light on some relevant aspects of a program of this nature. This can help to understand its importance in the context of educational policies and give space for comparative studies related to other similar programs. The selected criteria can serve as a starting point for further research, offering a more adjusted possibility of analysis to assessment policies, especially literacy – an effort that has not been found in the national literature on educational public policies.

Regarding the considerations that go beyond Paebes Alfa, attention is drawn to the mismatch between the educational literature's lack of interest in literacy assessment programs and the importance they assume in the context of state educational policy. Such importance is attested not only by the longevity of programs of this nature, as is the case with Paebes Alfa, but also by the managers who oversaw the program (political importance for management decision-making) and the specialists from research institutions that work in partnership with the Education Department (technical importance, with the program being an example and basis for the creation of new initiatives of the same nature in other contexts). At this point, it is worth suggesting that the observed mismatch may be the result of: a) The lack of interest in educational research on the view of managers regarding the questioned educational policies, as is the case of educational assessment, in a perspective other than disqualifying beforehand both the policy and its enforcer; b) the resistance to the topic of educational assessment; c) the absence of a national assessment of consolidated literacy, given the recent and discontinuous experience with ANA. All these elements make room for a new research effort to better understand how each one of them contributes to the near invisibility of literacy assessment programs in national education research.

REFERENCES

- BARBOSA, Liliane Cecília de Miranda; VIEIRA, Lívia Fraga. Avaliações externas estaduais: possíveis implicações para o trabalho docente. *Revista e-Curriculum*, São Paulo, v. 2, n. 11, p. 409-433, ago. 2013. Available at: <https://revistas.pucsp.br/index.php/curriculum/article/view/16615/12466>. Accessed on: Apr. 17, 2020.
- BONAMINO, Alicia; SOUSA, Sandra Zákia. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. *Educação e Pesquisa*, São Paulo, v. 38, n. 2, p. 373-388, abr./jun. 2012. Available at: <https://www.scielo.br/j/ep/a/rtQkYDSjky4mXG9TCrgRSqJ/?format=pdf&lang=pt>. Accessed on: July 2, 2020.
- BRASIL. Presidência da República. *Decreto Nº 6.094, de 24 de abril de 2007*. Dispõe sobre a implementação do Plano de Metas Compromisso Todos pela Educação, pela União Federal [...]. Brasília, DF: Presidência da República, 2007. Available at: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/decreto/d6094.htm. Accessed on: July 8, 2021.
- BROOKE, Nigel. High-stakes accountability, using teacher salary incentives in Brazil: an update. *Profesorado: Revista de currículum y formación del profesorado*, Granada, v. 20, n. 3, p. 207-250, 2016. Available at: <https://www.redalyc.org/pdf/567/56749100008.pdf>. Accessed on: Aug. 3, 2020.
- CALDERÓN, Adolfo Ignacio. Usos e apropriações das avaliações em larga escala. In: QUIOSSA, Amanda Sangy *et al.* (org.). Planos de Ação para a rede estadual de ensino de Minas Gerais. Editora CAEd/UFJF: Juiz de Fora, 2017. p. 31-45. (Série Diálogos e Proposições, v. 2).
- CASTRO, Maria Helena Guimarães. A consolidação da política de avaliação da educação básica no Brasil. *Meta: avaliação*, Rio de Janeiro, v. 1, n. 3, p. 271-296, set./dez. 2009. Available at: <https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/51>. Accessed on: June 12, 2020.
- CÔCO, Dilza. *Avaliação externa da alfabetização: o PAEBES-ALFA no Espírito Santo*. 2014. 404 f. Tese (Doutorado em Educação) – Universidade Federal do Espírito Santo, Vitória, 2014.
- DEUBEL, André-Nöel Roht. *Políticas públicas: formulación, implementación y evaluación*. Bogotá: Aurora, 2009.
- DRAIBE, Sônia Miriam. Avaliação de implementação: esboço de uma metodologia de trabalho em políticas públicas. In: BARREIRA, Maria Cecília Roxo Nobre; CARVALHO, Maria do Carmo Brandt (org.). *Tendências e perspectivas na avaliação de políticas e programas sociais*. São Paulo: IEE/PUC-SP, 2001. p. 13-42.
- ESPÍRITO SANTO. Lei complementar N. 504, de 23 de novembro de 2009. Institui a Bonificação por Desempenho, no âmbito da Secretaria de Estado da Educação - SEDU,

e dá outras providências. *Diário Oficial dos Poderes do Estado*: Vitória, ES, n. 22.459, p. 1, 23 ago. 2009.

ESPÍRITO SANTO. Secretaria de Estado da Educação. *Revista Pedagógica – Língua Portuguesa e Matemática, 3º ano do ensino fundamental*. Vitória: SME, 2017. Available at: <https://avaliacaoemonitoramentoespíritosanto.caeddigital.net/#!/colecões>. Accessed on: May 15, 2020.

ESPÍRITO SANTO. Governo do Espírito Santo. *Portal da Transparência do Poder Executivo do Espírito Santo*, c2019. Página inicial. Available at: <https://transparencia.es.gov.br>. Accessed on: 16 jun. 2019.

ESTEBAN, Maria Teresa. Considerações sobre a política de avaliação da alfabetização: pensando a partir do cotidiano escolar. *Revista Brasileira de Educação*, Rio de Janeiro, v. 17, n. 51, p. 573-592, set./dez. 2012. Available at: <https://doi.org/10.1590/S1413-24782012000300005>. Accessed on: July 19, 2020.

FREITAG, Raquel Meister Ko; ALMEIDA, Ayane Nazarela Santos; ROSÁRIO, Mônica Maria Soares. Contribuições para o aprimoramento da Provinha Brasil enquanto instrumento diagnóstico do nível de alfabetização e letramento nas séries iniciais. *Revista Brasileira de Estudos Pedagógicos*, Brasília, v. 94, n. 237, p. 390-416, maio/ago. 2013. Available at: <https://www.scielo.br/j/rbeped/a/dW6xSbk6RCgrMXMR7PqfPWd/abstract/?lang=pt>. Accessed on: July 20, 2020.

FREY, Klaus. Políticas públicas: um debate conceitual e reflexões referentes à prática da análise de políticas públicas no Brasil. *Revista Planejamento e políticas públicas*, Brasília, n. 21, p. 212-259, jun. 2000. Available at: <https://www.ipea.gov.br/ppp/index.php/PPP/article/view/89/0>. Accessed on: June 15, 2020.

GONTIJO, Cláudia Maria Mendes. *Alfabetização: políticas mundiais e movimentos nacionais*. Campinas: Autores Associados, 2014.

JANNUZZI, Paulo de Martino. A implementação no centro da Avaliação de Políticas Públicas. *Revista Aval*, Fortaleza, v. 2, n. 16, p. 64-80. jul./dez. 2019. Available at: http://repositorio.ufc.br/bitstream/riufc/49662/1/2019_art_pmjannuzzi.pdf. Accessed on: May 13, 2020.

LEAL, Telma Ferraz *et al.* Habilidades de compreensão leitora: seu ensino e sua avaliação pela Provinha Brasil. *Revista Brasileira de Educação*, Rio de Janeiro, v. 22, n. 68, p. 187-211. jan./mar. 2017. Available at: <https://doi.org/10.1590/S1413-24782017226810>. Accessed on: Apr. 14, 2020.

LIMA, Naira da Costa Muylaert. *Educação e desigualdades sociais: suas relações no estado do Espírito Santo*. 2016. 238 f. Tese (Doutorado em Educação) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2016.

MAIA, Cláudio Machado; GRANDO, Johnny Luiz. Metodologia para definição de tipologias de políticas públicas, para entender a elaboração, implementação e avaliação.

Colóquio: *Revista do Desenvolvimento Regional*, Taquara, v. 13, n. 2, p. 159-178. jul./dez. 2016. Available at: <https://doi.org/10.26767/coloquio.v13i2.497>. Accessed on: Apr. 15, 2020.

MARTINS, Ângela Maria. O campo das políticas públicas de educação: uma revisão de literatura. *Est. Aval. Educ.*, São Paulo, v. 24, n. 56, p. 276-299, set./dez. 2013. Available at: <https://doi.org/10.18222/eaec245620132741>. Accessed on: Apr. 15, 2020.

MORAIS, Artur Gomes de. Políticas de avaliação da alfabetização: discutindo a Provinha Brasil. *Revista Brasileira de Educação*, Rio de Janeiro, v. 17, n. 51, p. 551-572, set./dez. 2012. Available at: <https://doi.org/10.1590/S1413-24782012000300004>. Accessed on: Apr. 15, 2020.

OLIVEIRA, Dayane Louzada de. *Análise dos resultados do PAEBES de Ciências da Natureza do município de Marataízes-ES e seu reflexo na prática docente*. 2018. 111 f. – Dissertação (Mestrado em Ensino, Educação Básica e Formação de Professores) – Universidade Federal do Espírito Santo, Alegre, 2018.

OLIVEIRA JÚNIOR, Ruy Barboza de. *As políticas educacionais do segundo mandato do governo Paulo Hartung e seus impactos no trabalho docente*. 2013. 148 f. Dissertação (Mestrado em Educação) – Programa de Pós-Graduação em Educação, Universidade Federal do Espírito Santo, Vitória, 2013.

PAEBES. *Portal do Programa Avaliação e Monitoramento da Educação Básica*, 2020. Available at: <https://avaliacaoemonitoramentooespiritosanto.caedigital.net/#!/pagina-inicial>. Accessed on: July 20, 2020.

PASQUALI, Luiz. *Psicometria*. Teoria dos testes na Psicologia e na Educação. 5. ed. Petrópolis: Vozes, 2013.

PAULA, Patrícia de Lima. *A influência da elaboração de itens na eficácia das avaliações em larga escala*. 2018. 100 f. Dissertação (Mestrado em Gestão e Avaliação da Educação Pública) – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora, Juiz de Fora, 2018.

PEREIRA, Selma Lúcia de Assis. *PAEBES: modos, formas e diálogos a partir dos usos dos resultados em língua portuguesa da avaliação externa estadual no município de Cariacica-ES*. 2014. 233 f. Tese (Doutorado em Educação) – Universidade Federal do Espírito Santo, Vitória, 2014.

PORTELA, Máira Miranda. *Alfabetização matemática: um paralelo entre a avaliação nacional e cenários estaduais*. 2016. 118 f. Dissertação (Mestrado em Gestão e Avaliação da Educação Pública) – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora, Juiz de Fora, 2016.

REZENDE, Wagner Silveira. *Avaliação em larga escala da alfabetização: os casos de Paebes Alfa, Proalfa e Spaece Alfa*. 2020. 366 f. Tese (Doutorado em Educação) – Programa de Pós-Graduação em Educação, Universidade Federal de Juiz de Fora, Juiz de Fora, 2020.

ROCHA, Gladys; MARTINS, Raquel Fontes. Construção de um corpus de escrita infantil com itens de avaliações. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 51, p. 82-100, jan./abr. 2012. Available at: <https://doi.org/10.18222/eaec235120121949>. Accessed on: May 20, 2020.

SEGATTO, Catarina Ianni. *Como ideias se transformam em reformas: um estudo comparativo das mudanças educacionais orientadas pelo desempenho nos estados brasileiros*. 2011. Dissertação (Mestrado em Administração Pública) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2011.

SILVA, Josiane Toledo Ferreira. *A escrita na avaliação da alfabetização em larga escala*. 2013. 136 f. Tese (Doutorado em Educação) – Programa de Pós-Graduação em Educação, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013.

SOARES, Magda. Letramento e alfabetização: as muitas facetas. *Revista Brasileira de Educação*, Campinas, n. 25, p. 5-17, 2004. Available at: <https://doi.org/10.1590/S1413-24782004000100002>. Accessed on: May 20, 2020.

SOARES, Magda. *Alfabetização e letramento*. 5. ed. São Paulo: Contexto, 2007.

Text received on 02/07/2021.

Text approved on 05/03/2021.