Psychology and Education

PTP
Psicologia
Teoria e Prática

Standardization of the Instrument for Brief Assessment of Reading, Writing, and Comprehension − IBALEC

Gislaine Gasparin Nobile[1], Flávio Rebustini[2], e Sylvia Domingos Barrera[1]

[1] University of São Paulo (USP), Ribeirão Preto, SP, Brazil
[2] University of São Paulo (USP), São Paulo, SP, Brazil

Author Note

Gislaine G. Nobile (iD) https://orcid.org/0000-0001-5463-4779

Flávio Rebustini (iD) https://orcid.org/0000-0002-3746-3266

Sylvia D. Barrera (iD) https://orcid.org/0000-0001-7924-2755

Correspondence concerning this article should be addressed to Gislaine Gasparin Nobile, University of São Paulo, Faculty of Philosophy, Sciences, and Letters of Ribeirão Preto, Postgraduate Program in Psychology − University of São Paulo. Avenida Bandeirantes, 3900, Monte Alegre, 14040-901, Ribeirão Preto-SP, Brazil. Email: gislainenobile@yahoo.com.br

Conflict of Interest: None declared.

Abstract

Grounded in cognitive psychology, the IBALEC is designed to assess basic reading, writing and reading comprehension skills among 1st–5th year students in the first cycle of elementary education in Brazil. This study presents the results of the instrument standardization process, in which the IBALEC was individually administered to 825 public school students (439 boys; 386 girls; median age = 8.5 years; SD = 1.65). The results show that overall scores increase as the academic years progress, demonstrating that the instrument has good sensitivity for differentiating between different levels of literacy skills across academic years. The standardization of scores into three performance bands (lag in performance, satisfactory performance and high performance) for each academic year meets the instrument's objective of offering health and education professionals a brief instrument for assessing basic aspects of mastery of written language.

Keywords: psychometrics, educational assessment, reading, literacy, comprehension

NORMATIZAÇÃO DO INSTRUMENTO PARA BREVE AVALIAÇÃO DA LEITURA, ESCRITA E COMPREENSÃO – IBALEC

Resumo

Elaborado com base no referencial teórico da Psicologia Cognitiva, o IBALEC destina-se a avaliar as habilidades básicas de leitura, escrita e compreensão da leitura em alunos do 1º ao 5º ano do Ensino Fundamental. Neste estudo são apresentados os resultados das análises de normatização do Instrumento, que foi aplicado individualmente em 825 alunos (439 do sexo masculino e 386, feminino; M=8,5 anos; DP=1,65) da rede pública de ensino. Os resultados indicam um aumento do escore total obtido pelos participantes de acordo com o aumento dos anos escolares, o que demonstra boa sensibilidade do Instrumento na discriminação das habilidades avaliadas em função da escolaridade. A normatização dos resultados, estabelecendo três faixas de desempenho (defasado, satisfatório e superior) para cada ano escolar mostrou-se adequada aos objetivos do IBALEC, de oferecer a profissionais da saúde e educação uma ferramenta que permita a avaliação breve de aspectos básicos do domínio da língua escrita.

Palavras-chave: psicometria, avaliação educacional, leitura, alfabetização, compreensão

ESTANDARIZACIÓN DEL INSTRUMENTO DE EVALUACIÓN BREVE DE LECTURA, ESCRITURA Y COMPRENSIÓN – IBALEC

Resumen

Basado en el marco teórico de la Psicología Cognitiva, el IBALEC pretende evaluar las habilidades básicas de lectura, escritura y comprensión lectora en estudiantes – 1° a 5° año – Enseñanza Fundamental. Este estudio presenta los resultados del análisis de estandarización del Instrumento, aplicado a 825 alumnos (439 varones; 386 mujeres; M=8,5 años; DT=1,65) del sistema escolar público. Los resultados indican un aumento en el puntaje total obtenido por los participantes de acuerdo con el aumento de años escolares, lo que demuestra una buena sensibilidad del Instrumento en la discriminación de las habilidades evaluadas en función de la escolaridad. La estandarización de los resultados, estableciendo tres rangos de desempeño (rezagado, satisfactorio y superior) para cada año escolar se mostró adecuada a los objetivos del IBALEC, ofrecer a los profesionales de la salud y la educación una herramienta que permita la evaluación breve de aspectos básicos del dominio del lenguaje escrito.

Palabras clave: psicometría, evaluación educacional, lectura, alfabetización, comprensión

Unlike oral language learning, which originates from biological processes activated by the child's contact with his/her social group, written language learning involving alphabetic systems depends on explicit teaching, requiring a high level of abstraction, elaboration and control, primarily due to the arbitrary nature of the representation of speech sounds using graphic characters (Maluf & Gombert, 2008).

According to the guidelines of the National Assessment of Literacy (ANA), the literacy development cycle includes the first three years of elementary education and "should provide students more than the simple acquisition of a code, prompting them to construct meaning and written language use in different contexts of communication" (Oliveira et al., 2022, p. 340). However, given the complexity of learning the writing system and reading processes, including decoding skills and the more complex skills of reading comprehension and text production, many students demonstrate difficulties during the literacy development process.

Indeed, poor academic performance is a concern both among professionals in the field of education and health and the affected children's parents and guardians. When detected early on, the damage caused by poor academic performance can be mitigated, avoiding future problems in different contexts of professional and personal life (Rosa et al., 2022).

It is important, however, to make a distinction between the terms learning "difficulty" and learning "disorder", since each one implies different causes and, very often, distinct approaches to addressing the problem. Learning difficulties are generally related to environmental factors or aspects that are extrinsic to the child, such as inappropriate pedagogical approaches and/or adverse social, cultural and family conditions. According to Rotta (2016), a normal brain structure and adequate functional and neurochemical conditions do not guarantee proper learning. The act of learning is an act of brain plasticity, meaning that "learning difficulties is a generic term encompassing a heterogenous group of problems that are capable of altering the child's learning possibilities, irrespective of his/her neurological conditions" (p. 97-98).

In contrast, according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR, 2023), learning disorders are caused by factors that are intrinsic to the student, implying biological origins, including an interaction of genetic, epigenetic, and environmental factors, which affect cognitive abilities, leading to persistent difficulties with academic learning.

Capellini and Conrado (2009) suggest that students with literacy difficulties are often confused with children with learning disorders. According to Zorzi (2003), many children from both public and private schools have experienced school failure due to difficulties with learning to read and write, leading the author to conclude that many educational approaches may not be meeting the specific needs of these students. This has been exacerbated by two years of pandemic, during which many students had remote classes while others, for various reasons, did not have classes at all. Recent studies have shown the harmful effects of social isolation during the Covid-19 pandemic on learning early reading and writing skills, especially among public school students, who had greater difficulty keeping up with remote classes and school activities (Bartholo et al., 2022).

Children with learning difficulties are referred to health professionals (pediatricians, neurologists and psychologists). However, these professionals do not generally have the brief instruments necessary to assess the degree of educational lag, especially with regard to reading and writing. This type of assessment is crucial to provide insights into the severity of the problem and criteria for diagnosis and future referral.

The instruments available in Brazil generally assess word reading and writing skills in a more systematic and often exhaustive manner, leaving aside the more complex skills involved in writing and comprehending sentences and/or texts, as is the case with the School Achievement Test II – TDE (Stein et al., 2019), Single Word and Pseudo-word Reading Task – LPI (Salles et al., 2017) and Dictation Writing Test (Seabra et al., 2013).

The Instrument for Brief Assessment of Reading, Writing, and Comprehension (*Instrumento para Breve Avaliação da Leitura, Escrita e Compreensão* or IBALEC, in Portuguese) is designed to evaluate basic reading, writing and reading comprehension skills among students in the first cycle of elementary education. Grounded in cognitive psychology, the individually-administered IBALEC consists of six tasks (or subtests): in the first task, the child is asked to write the letters of the alphabet in their conventional order or the letters that he/she knows/remembers. The test administrator then asks the child to name the letters he/she wrote from last to first to avoid mechanical repetition. Points are only given for correctly written and read letters. This task is warranted by studies showing that letter knowledge is a key predictor of reading and writing acquisition in alphabetic languages (Justi et al., 2020; Morais et al., 2013). However, the task should only be applied when the child completely fails the word reading and writing tests (tasks 2 and 3). In such cases the application of tasks 4, 5 and 6 is not recommended.

In the second task the child is asked to read 13 regular and irregularly spelled words arranged in order of syllabic structure complexity, from simple (consonant-vowel – CV), such as *SALA* (room) and *BIFE* (steak), to more complex (CCV, CVC), such as *CHUVA* (rain) and *CONFUSÃO* (confusion). For the third task the child is asked to write the words corresponding to seven pictures of animals whose names are made up of simple syllables (CV), and complex syllables (CVC, CCV, CVV), such as *FORMIGA* (ant), *COBRA* (snake) and *PEIXE* (fish).

These tasks were designed based on Ehri's (1998; 2013) four-phase model of word reading and writing development, which classifies each read or written word into one of the following phases: pre-alphabetic, partial alphabetic, full alphabetic and consolidated alphabetic. These phases indicate the child's level of knowledge, from the ability to make connections between graphemes and phonemes to the extent to which these connections become automatic and embedded in the long-term memory. It is important to emphasize that these phases are not totally symmetrical, as the child may progress more rapidly with writing than reading or vice versa, demonstrating a certain mismatch between strategies used during the literacy process (Ehri, 2013).

In the fourth task, the child is asked to write five sentences based on five visual stimuli (*palhaço* – clown; *chapéu* – hat; *camisa* – shirt; *elefante* – elephant; and *escorregador* - slide) after

providing an example sentence using a picture of a giraffe. The aim is to test the syntactic and semantic adequacy of the sentences produced and adherence to writing conventions, such as the correct use of upper and lower case letters and punctuation.

The fifth and sixth tasks test different levels of complexity of reading comprehension. In the fifth task the child is asked to read three sentences, each followed by three pictures. For each sentence the child must choose the picture that best represents its meaning. This task is directed mainly at children in the first years of elementary education, who often do not have an automatic mastery of the writing system but have basic decoding skills, meaning they are able to understand short simple sentences. The sixth task, which is also a reading comprehension test, consists of a simple narrative text (89 words) with six questions requiring a written answer aimed at testing information retrieval and ability to draw inferences.

The assessment of reading comprehension skills was based on the approach proposed by Perfetti, Landi and Oakhill (2013), which suggests that comprehension processes occur at multiple levels: word level (lexical processes), sentence level (syntactic processes), and text level. Another critical process for producing higher-level comprehension is drawing inferences. The language of any text, spoken or written, is not completely explicit, thus requiring the reader to connect elements of text with their knowledge of the world to make it coherent by filling the "gaps" left by the author.

Previous studies investigating the content and construct validity (internal structure) of the IBALEC have been published recently (Nobile et al., 2021; 2023). The aim of this study was to standardize the instrument.

Method

Participants

The study sample comprised 825 1st-5th year students from two public elementary schools in a large city in the state of São Paulo selected through convenience sampling and distributed as follows: 160 5th-year students (90 boys; 70 girls; median age = 10.7 years); 161 4th-year students (85 boys; 76 girls; median age = 9.7 years); 165 3rd-year students (85 boys; 80 girls; median age = 8.5 years); 178 2nd-year students (97 boys; 81 girls; median age = 7.3 years); and 161 1st-year students (82 boys; 79 girls; median age = 6.4 years). No inclusion or exclusion criteria were used. Only students who did not return the signed informed consent form did not participate in the study (a total of 6 students).

Instrument

The Instrument for the Brief Assessment of Reading, Writing, and Comprehension (IBALEC) was used to assess performance on basic reading, writing and reading comprehension skills among students in the first cycle of elementary education. Grounded in cognitive psychology, the individually-administered IBALEC consists of six tasks (or subtests) described in the introduction.

Procedures

The study protocol was approved by the research ethics committee at the institution to which the study is affiliated (CAAE: 02902118.2.0000.5407). The study was then presented in the participating schools. The children's parents/guardians signed a parental informed consent form and child assent was obtained from the participants. The instrument was administered individually during normal class hours in a room at the school arranged by the school managers. Each test lasted between 10 and 30 minutes. The data were collected during the period August–December 2019, before the onset of the pandemic.

Data analysis – procedures for the standardization of the IBALEC

First, we performed an exploratory descriptive analysis of the overall and domain scores obtained by each academic year. The results were presented as medians, interquartile ranges, ranges, minimums, maximums, means and standard deviations. The IBALEC consists of 5 domains with a maximum score of 121: 52 points for word reading; 28 points for word writing; 20 points for sentence writing; 3 points for sentence comprehension; and 18 points for text comprehension.

The IBALEC standardization process consisted of three stages. In the first stage, three classification options were defined for each literacy skill domain: 3 bands; 4 bands; and 5 bands. The 3 band option was defined based on the distribution of the participants using the 1st quartile, the interquartile range and 3rd quartile. The 4 band system was defined using 4 quartiles. In the 5 band system, each band consisted of 20 percentiles. Testing three types of score distribution enabled us to identify the most effective student classification option.

The second stage consisted of verifying and confirming the most effective and accurate student classification option. Two-step cluster analysis was used to determine which of the three options ensured the correct positioning of participants in the bands. As Hair et al. (2019, p. 198) highlight, "cluster analysis groups individuals or objects into clusters so that objects in the same cluster are more similar to one another than they are to objects in other clusters", adding that "The attempt is to maximize the homogeneity of objects within the clusters while also maximizing the heterogeneity between the clusters". The most accurate classification option was extrapolated to the 5 domains of the IBALEC. These procedures were adopted for each academic year.

In the third stage, multivariate analysis was performed to determine whether the instrument differentiated academic years according to their scores, i.e. whether the instrument has the capacity to demonstrate student development based on their scores, as well as identify students with a lag in performance, moderate performance and high performance in each academic year. Since the scores are ordinal data, it was necessary to use non-parametric techniques. We therefore applied the Kruskal–Wallis test with Dunn's post hoc test, Bonferroni correction for Type I errors and η^2 H effect size.

All crude "p-values" except those under 0.001 are reported, as recommended by Wasserstein and Lazar (2016). This procedure aims to avoid inappropriate p-value interpretation

practices (Friese & Frankenbach, 2020). P-value levels are therefore treated as similar or not to the variable scores and as "a violation of the null hypothesis" (VNH) or "without violation of the null hypothesis (WVNH)" (Wasserstein et.al., 2019). All analyses included estimates of effect size, as recommended by the literature (Wasserstein et al., 2019; Tomczak & Tomczak, 2014).

Results

The descriptive statistics for the overall and domain scores obtained by each academic year are shown in Table 1. The results clearly show that the median overall score increases as the academic years progress. It is also important to note that in each academic year the scores achieved by the students encompassed the entire spectrum of possible scores in practically every domain, indicating low likelihood of bias and showing that the instrument has the sensitivity to capture the full spectrum of levels assessed by the instrument.

Table 1

Descriptive statistics: scores by academic year

Year	Domain	N	Md	IQR	Range	Min	Max	Mean	SD
1st	Word reading	161	13.00	21.00	52	0	52	20.81	15.92
	Word writing		13.00	21.00	28	0	28	12.58	10.10
	Sentence writing		0.00	0.00	11	0	11	1.39	3.00
	Sentence comprehension		0.00	2.00	3	0	3	0.77	1.26
	Text comp		0.00	0.00	13	0	13	1.45	3.40
	Total		26.00	44.00	102	0	102	37.00	30.98
2nd	Word reading	178	46.00	32.00	52	0	52	36.50	17.40
	Word writing		25.00	10.00	28	0	28	20.53	8.81
	Sentence writing		5.50	10.00	18	0	18	2.58	5.21
	Sentence comprehension		3.00	3.00	3	0	3	1.80	1.38
	Text comp		9.00	12.00	17	0	17	7.05	5.63
	Total		90.00	66.50	111	0	111	71.46	35.77
3rd	Word reading	165	50.00	5.00	52	0	52	45.13	12.18
	Word writing		26.00	3.00	28	0	28	24.14	5.56
	Sentence writing		9.00	7.00	20	0	20	8.81	5.4
	Sentence comprehension		3.00	1.00	3	0	3	2.42	1.06
	Text comp		12.00	6.00	17	0	17	10.04	5.05
	Total		101.00	20.00	116	0	116	90.50	26.61

Table 1

Descriptive statistics: scores by academic year

Year	Domain	N	Md	IQR	Range	Min	Max	Mean	SD
4th	Word reading	161	52.00	2.00	52	0	52	49.03	7.37
	Word writing		27.00	1.00	28	0	28	25.85	3.53
	Sentence writing		12.00	7.00	20	0	20	10.70	4.93
	Sentence comprehension		3.00	0.00	3	0	3	2.76	0.66
	Text comp		13.00	2.00	17	0	17	12.17	3.27
	Total		106.00	12.00	106	13	119	100.50	16.24
5th	Word reading	160	52.00	2.00	39	13	52	50.44	5.20
	Word writing		27.00	1.00	28	0	28	25.97	3.57
	Sentence writing		12.00	5.25	20	0	20	11.55	4.47
	Sentence comprehension		3.00	0.00	3	0	3	2.85	0.51
	Text comprehension		13.00	2.25	17	0	17	12.87	2.50
	Total		107.00	8.00	104	13	117	103.70	13.73

N – Number of participants; Md – median; IQR – interquartile range; Min – minimum; Max – maximum; SD – standard deviation.

The results of the cluster analysis (Table 2) suggest that the three-band student classification option was more accurate for four of the five academic years (1st, 2nd, 3rd and 5th), correctly classifying 100% of the scores. The five-band classification option was the most accurate for the 4th year. Since the predictive value of the three-band classification was also high (0.87) for the 4th year, this option was deemed to have the highest capacity to correctly classify the students over all academic years. The three bands were characterized as follows: Band 1 – lag in performance; Band 2 – satisfactory performance; and Band 3 – high performance.

Table 2

Cluster analysis – predictive value of the classification options

		Option		
		3-band	4-band	5-band
Overall score	1st year	1.00	0.96	0.92
	2nd year	1.00	0.97	0.73
	3rd year	1.00	0.97	0.70
	4th year	0.87	0.93	1.00
	5th year	1.00	0.97	0.63

Table 3 shows the band cut-off scores for each domain in each academic year. For most of the domains and years it was possible to maintain the three-band classification. However, in some cases the scores exhibited very high homogeneity, resulting in categories with merged bands. In the 1st year, this was the case with sentence writing, sentence comprehension and text comprehension, where it was only possible to classify the scores in two bands (high performance and the merging of bands 1 and 2).

This was the case with the domain sentence comprehension in the 3rd, 4th and 5th years, in which scores under 3 were deemed to indicate lag in performance. For the 5th year, bands 2 and 3 (satisfactory performance and high performance) overlapped in the domain word writing, and band 1 was maintained.

Table 3

Classification by domain and academic year

		Band 1 Lag in performance	Band 2 Satisfactory performance	Band 3 High performance
1st year	Word reading (0–52)	≤ 12	13–32	≥ 33
	Word writing (0 –28)	≤ 2	3–22	≥ 23
	Sentence writing (0–20)	≤ 1		> 1
	Sentence comprehension (0–3)	≤ 2		> 2
	Text comprehension (0–18)	≤ 1		> 1
	Overall score (0–121)	13	14–56	≥ 57
2nd year	Word reading	≤ 19	20–50	≥ 51
	Word writing	≤ 17	18–26	≥ 27
	Sentence writing	0	1–9	≥ 10
	Sentence comprehension	0	1–2	≥ 3
	Text comprehension	0	1–11	≥ 12
	Overall score	≤ 34	35–100	≥ 101
3rd year	Word reading	≤ 47	48–51	≥ 52
	Word writing	≤ 24	25–26	≥ 27
	Sentence writing	≤ 6	7–12	≥ 13
	Sentence comprehension	≤ 2		3
	Text comprehension	≤ 8	9–13	≥ 14
	Overall score	≤ 87	88–106	≥ 107
4th year	Word reading	≤ 50	51	≥ 52
	Word writing	≤ 26	≥ 27	
	Sentence writing	≤ 7	8–13	≥ 14
	Sentence comprehension	≤ 2		3
	Text comprehension	≤ 12	13	≥ 14
	Overall score	≤ 97	98–108	≥ 109

Table 3

Classification by domain and academic year

		Band 1 Lag in performance	Band 2 Satisfactory performance	Band 3 High performance
5th year	Word reading	≤ 50	51	≥ 52
	Word writing	≤ 26	≥ 27	
	Sentence writing	≤ 9	10–13	≥ 14
	Sentence comprehension	≤ 2		3
	Text comprehension	≤ 12	13	≥ 14
	Overall score	≤ 102	103–109	≥ 110

The multivariate analysis of the overall score resulted in $H_{(4)}$ = 376.721; p < 0.001 and η^2_H = 0.46 (large effect size), indicating a difference in variance of scores as the academic years progress. The results of Dunn's post hoc test with Bonferroni correction suggest a violation of the null hypothesis (p < 0.001) for all interactions between academic years, with one exception: the overall scores for the 4th and 5th years did not show any violation of the null hypothesis (p = 0.445). As a reference point, the median overall score of the 4th year was 100.50 (SD = 16.24), while that of the 5th year was 103.70 (SD = 13,73), indicating greater dispersion of results for the 4th year. In any case, it is evident that there are important variations in scores and that scores increase as the academic years progress.

Word reading domain – the multivariate analysis resulted in $H_{(4)}$ = 377.810; p < 0.001 and η^2_H = 0.42 (large size effect), indicating a difference in variance of scores as the academic years progress. The results of the post hoc test suggest a violation of the null hypothesis (p < 0.001) for all interactions between academic years, with two exceptions: the interaction between the 3rd and 4th years was weak (p = 0.017) and the scores for the interaction between the 4th and 5th years once again did not show any violation of the null hypothesis (p = 0.262).

Word writing domain – the multivariate analysis resulted in $H_{(4)}$ = 269.650; p < 0.001 and η^2_H = 0.33 (medium to large size effect). The results of the post hoc test suggest a violation of the null hypothesis (p < 0.001) for all interactions between academic years, with the same two exceptions found in the word reading domain: the violation for the interaction between the 3rd and 4th years was insignificant (p = 0.006) and the scores for the interaction between the 4th and 5th years once again did not show any violation of the null hypothesis (p = 1.000). These p-values are explained by the range of scores, which was the same in both years, and similarity between the mean of the values and deviation (4th year, mean = 25.85, deviation = 3.53; 5th year, mean = 25.97, deviation = 3.57).

Sentence writing domain – the multivariate analysis resulted in $H_{(4)}$ = 312.206; p < 0.001 and η^2_H = 0.38 (medium to large effect size). The results of the post hoc test suggest a violation of the null hypothesis (p < 0.001) for all interactions between academic years, with two

exceptions: the violation in the interaction between the 3rd and 4th years was weak (p= 0.011) and once again the scores for the interaction between the 4th and 5th years did not show any violation of the null hypothesis (p = 1.000).

Sentence comprehension domain – the multivariate analysis resulted in $H_{(4)}$ = 265.925; p < 0.001 and η^2_H = 0.33 (medium to large effect size). The results of the post hoc test suggest a violation of the null hypothesis (p < 0.001) for all interactions between academic years, except between the 3rd and 4th and 3rd and 5th years (p = 0.008), and once again the scores for the interaction between the 4th and 5th years did not show any violation of the null hypothesis (p = 1.000).

Text comprehension domain – the multivariate analysis resulted in $H_{(4)}$ = 265.119; p < 0.001 and η^2_H = 0.45 (large effect size). The results of the post hoc test suggest a violation of the null hypothesis (p < 0.001) for all interactions between academic years, with three exceptions: the violation for the interaction between the 3rd and 4th and 3rd and 5th years (p = 0.008) was weak and once again the scores for the interaction between the 4th and 5th years did not show any violation of the null hypothesis (p = 1.000).

A final test was performed to verify whether the schools where data collection was carried out could bring about bias in the results. The results of the Mann–Whitney test (U= 74123.00; p = 0.76) showed that the scores obtained using the instrument were similar in both schools, including an irrelevant effect size (r_s = – 0.01).

Discussion

The findings show that overall scores increase as the academic years progress, confirming that the IBALEC has good sensitivity for the assessment of reading and writing skills at the different levels of complexity tested by the instrument. The standardization of scores into three performance bands (lag in performance, satisfactory performance and high performance) for each academic year seems to meet the instrument's objective, which is to offer health and education professionals a brief instrument for assessing basic aspects of mastery of written language, based on a robust sample of 1st–5th year elementary school students.

However, the findings reveal that some domains in some academic years exhibited homogeneity of performance across bands, such as the 1st year, in which scores for more complex skills like sentence writing and comprehension did not differentiate students with a lag in performance from those with satisfactory performance. This means that lack of success in these tasks does not yet constitute a lag in learning but is rather expected (satisfactory) at this academic level. This may be explained by the fact that these children are in the early years of formal schooling and that the sample is made up exclusively of public school students, whose performance in reading and writing tasks tends to be behind those studying at private schools (Rosa et al., 2022). However, the same lack of differentiation, this time in the opposite direction, was observed in older students (4th and 5th years) for more basic skills such as word writing, in which the instrument does not differentiate between satisfactory and high performance. This

appears to reveal a "ceiling effect", in which scores indicating satisfactory performance overlap or are very close to the maximum possible score for the domain. However, overall IBALEC scores across all domains appear to differentiate well between academic years.

The findings also show that scores differ substantially between the first three years and between these years and the 4th and 5th years. The variations in scores between 4th and 5th year students were smaller, with no substantial differences. This is to be expected since 4th and 5th year students are expected to have adequate literacy levels, with the national curriculum (BNCC, Brazil, 2019) and National Pact for Literacy at the Right Age (PNAIC, Brazil, 2014) stating , respectively, that this stage of learning is expected to be completed by the end of the 2nd and 3rd years of elementary education. It is also important to highlight that the instrument did not fail to identify students in the last two academic years with lag in performance. The results therefore confirm that the IBALEC is an effective instrument for identifying reading and writing lag among 1st-5th year students in the first cycle of elementary education in Brazil, which is one of its primary objectives.

Final Considerations

The findings reveal that overall scores increased as the academic years progress and that the three-band classification system (band 1 – lag in performance; band 2 – satisfactory performance; and band 3 – high performance) was the most accurate option for differentiating student scores. The results also show that the instrument is suitable for use at a regional level and can contribute to the diagnosis of learning disorders and the implementation of interventions to tackle these problems. However, one of the limitations of this study is the fact that our sample was limited to a single region. For the instrument to be used at national level, the sample should therefore be increased to include other regions of the country. The Appendix presents the final version of the IBALEC and instructions on administration and scoring.

References

APA – Associação Psiquiátrica Americana (2023). *Manual Diagnóstico e Estatístico dos Transtornos Mentais – DSM-5-TR.* Artmed.

Bartholo, T. L., Koslinski, M. C., Tymms, P., & Castro, D. L. (2022). Learning loss and learning inequality during the Covid-19 pandemic. *Ensaio: aval.pol.públ. educ.*, 1-22. https://doi.org/10.1590/S0104-40362022003003776

Brasil – Pacto Nacional pela Alfabetização na Idade Certa – PNAIC (2014). *Currículo na perspectiva da inclusão e da diversidade: as Diretrizes Curriculares Nacionais da Educação Básica e o Ciclo de Alfabetização.* https://pacto.mec.gov.br/materiais-listagem/item/download/12_ab2c739d2e8293712078e7b6b0c12abb

Brasil. Ministério da Educação. Base Nacional Comum Curricular – BNCC. Versão final (2018). http://basenacionalcomum.mec.gov.br/wpcontent/uploads/2018/06/BNCC_EI_EF_110518_versaofinal_site.pdf

Bragança, A. D., & Capaneda, I. P. M. (1996). Alfabetização: Atividades. FTD.

Capellini, S. A., & Conrado, T. L. B. (2009). Desempenho de escolares com e sem dificuldades de aprendizagem de ensino particular em habilidade fonológica, nomeação rápida, leitura e escrita. *Rev. CEFAC*, *11*(2), 183-193. https://doi.org/10.1590/S1516-18462009005000002

Ehri, L. C. (1998). Learning to read and learning to spell are one and the same, almost. Em C. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: research, theory, and practice across languages* (pp.237-269). Lawrence Erlbaum Associates.

Ehri, L.C. (2013). O desenvolvimento da leitura imediata de palavras: fases e estudos. Em M. J. Snowling, & C. Hulme (Eds.), *A ciência da leitura* (pp.153-172). Penso.

Friese, M., & Frankenbach, J. (2020). p-Hacking and publication bias interact to distort meta-analytic effect size estimates. *Psychological Methods*, *25*(4), 456-471. https://doi.org/10.1037/met0000246

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R., & Tatham, R. L. (2019). *Multivariate Data Analysis.* (8ª ed). Cengage Learning.

Justi, C. N. G., Cunha, N., & Justi, F. R. (2020). Letter-name knowledge: Predicting reading and writing difficulties. *Estudos de Psicologia. (Campinas)*, *37*. https://doi.org/10.1590/1982-0275202037e180173

Maluf, M. R. & Gombert, J. E. (2008). Habilidades implícitas e controle cognitivo na aprendizagem da linguagem escrita. Em M. R. Maluf & S. R. K. Guimarães (Eds.), *Desenvolvimento da Linguagem Oral e Escrita* (pp.123-135). Editora UFPR.

Morais, J., Leite, I., & Kolinsky, R. (2013). Entre a pré-leitura e a leitura hábil: Condições e patamares da aprendizagem. Em M. R. Maluf, & C. Cardoso-Martins (Eds.), *Alfabetização no Século XXI: Como se aprende a ler e a escrever* (pp. 17-48). Penso.

Nobile, G. G., Barrera, S. D. & Rebustini, F. (2021). Avaliação da alfabetização: elaboração e validação de conteúdo do IBALEC. *Rev. Psicopedagogia*, *38*(117), 333-345. https://doi: 10.51207/2179-4057.20210028

Nobile, G. G., Rebustini, F., & Barrera, S. D. (2023). Análise de estrutura interna do IBALEC. *Psico-USF*, *28*(1), 103-116. https://doi.org/10.1590/1413-82712023280109

Oliveira, H. V., Pinho, D. M. V. & Senna, L. A. G. (2022). Políticas públicas na alfabetização: um diálogo com a avaliação nacional da alfabetização e o Programa Mais Alfabetização. *Ensaio: aval. pol. públ. Educ.*, *30*(115), 334-353. https://doi.org/10.1590/S0104-40362022003002686

Perfetti C. A., Landi N., & Oakhill J. (2013). A aquisição da habilidade de compreensão da leitura. Em M. J. Snowlin, & C. Hulme (Eds.). *A ciência da leitura* (pp 245-265). Penso.

Rosa, A. P., Santos, P. L., Taipe, C. M. M. & Dilleggi, E. S. (2022). Fatores de risco para baixo desempenho escolar: uma revisão integrativa. *Rev. Psicopedagogia*, *39*(120), 445-57. https://doi.org/10.51207/2179-4057.20220032

Rotta, N.T. (2016). Dificuldades para aprendizagem. Em N. T. Rotta, L. Ohlweiler, & R. Riesgo (Eds.), *Transtornos da Aprendizagem: abordagem neurobiológica e multidisciplinar* (pp. 94-104). Artmed.

Salles, J. F., Picollo, L. R., & Miná, C. S. (2017). LPI – *Avaliação de Leitura de Palavras e Pseudopalavras Isoladas* (Coleção ANELE 1). Vetor.

Seabra, A. G., Dias, N., & Capovilla, F. C. (2013). Prova de Escrita sob Ditado (versão reduzida). Em A. G. Seabra, & N. M. Dias (Eds.), *Avaliação neuropsicológica cognitiva: Leitura, escrita e aritmética* (pp. 54-59). Memnon.

Stein, L. M., Giacomoni, C. H., & Fonseca, R. P. (2019). *TDE II – Teste de Desempenho Escolar*. Vetor.

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in sport sciences*, *1*(21), 19-25. https://www.researchgate.net/publication/303919832_The_need_to_report_effect_size_estimates_revisited_An_overview_of_some_recommended_measures_of_effect_size

Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a world beyond "p<0,05". *The American Statistician*, *73*(1), 1-19. https://doi.org/10.1080/00031305.2019.1583913

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's Statement on p-Values: Context, Process, and Purpose. *The American Statistician*, *70*(2), 129-133. https://doi.org/10.1080/00031305.2016.1154108

Zorzi, J. L. (2003). *Aprendizagem e distúrbios da linguagem escrita: questões clínicas e educacionais*. Artmed.

<p style="text-align:center;color:red">APPENDIX</p>

INSTRUMENT FOR BRIEF ASSESSMENT OF READING, WRITING, AND COMPREHENSION – IBALEC

NAME: _____

DATE OF BIRTH: ___ /___ / ___   ACADEMIC YEAR: _____   DATE: ___ /___ / ___   AGE: _____

1– ALPHABET WRITING (OPTIONAL – can be used if the child fails the initial word reading and writing tasks)

_____

_____

2– WORD READING

BIFE – steak
SALA – room
POMADA – orchard
GELATINA – jelly
ROSA – rose
HOJE – today
VELHO – old
TÁXI – taxi
CHUVA – rain
ALFABETO – alphabet
CONFUSÃO – confusion
GALINHA – chicken
FLORESTA – forest
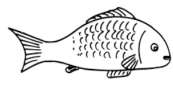
3– WORD WRITING

_____   _____

_____   _____

_____   _____

_____

4- SENTENCE WRITING

_____

_____

_____

_____

_____

_____
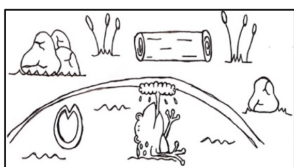
_____

_____

_____

_____
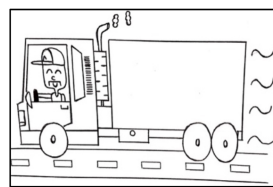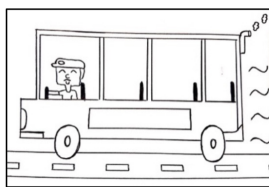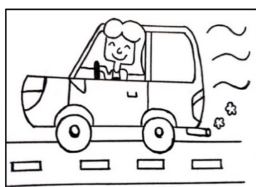
5– SENTENCE COMPREHENSION

O SAPO PULA NA LAGOA. (The frog jumps in the lake)
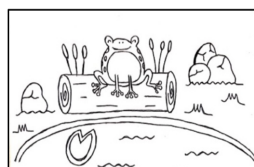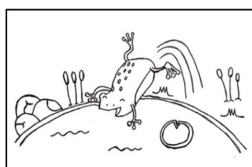
FUI NA FEIRA COM MINHA TIA E VI MUITAS FRUTAS E VERDURAS. (I went to the market with my aunt and saw lots of fruit and vegetables)

O MOTORISTA DESTE ÔNIBUS É O SENHOR JOÃO. (The driver of this bus is Mr. Jones)

Illustrator: Leonardo Ricieri Mantoani

6– TEXT COMPREHENSION

O DOMINGO ESTAVA TÃO QUENTE QUE O SENHOR ANTONIO TEVE A IDEIA DE LIGAR A MANGUEI-RA DE ÁGUA NO JARDIM PARA SEUS FILHOS SE MOLHAREM.

A MANGUEIRA COMEÇOU A GIRAR RÁPIDO, JOGANDO ÁGUA PARA TODO O LADO. MARCOS E ANA CORRIAM E PULAVAM NO JARDIM, DANDO TANTA GARGALHADA QUE CHAMOU A ATENÇÃO DOS AMIGUI-NHOS DA RUA. LOGO, TODA A GAROTADA ENTROU NO QUINTAL.

CANSADAS DE BRINCAR, SUJAS E MOLHADAS, AS CHILDREN DECIDIRAM TOMAR UM BANHO E DEPOIS DONA EMA AINDA SERVIU GUARANÁ GELADINHO PARA TODO MUNDO. FOI UMA FESTA!

(Sunday was so hot that Mr. Antonio had the idea of turning on the water hose in the garden so his children could get wet.

The hose started to spin fast and shoot water everywhere. Marcos and Ana ran and jumped around the garden, laughing so much that they caught the attention of their friends in the street. In no time at all, the kids entered the garden.

Tired of playing, dirty and wet, the children decided to have a shower and then Mrs. Ema served everyone cold soda. It was a party!)

Text adapted from Bragança and Carpaneda (1996).

1) *POR QUE O SENHOR ANTONIO LIGOU A MANGUEIRA DE ÁGUA NO JARDIM?* (Why did Mr. Antonio turn the water hose in the garden?)

_____
_____
_____

2) *POR QUE A GAROTADA ENTROU NO QUINTAL?* (Why did the kids enter the garden?)

_____
_____
_____

3) *POR QUE AS CRIANÇAS DECIDIRAM TOMAR UM BANHO?* (Why did the kids decide to have a shower?)

_____
_____
_____

4) *O QUE AS CRIANÇAS FIZERAM DEPOIS DO BANHO?* (What did the kids do after they had a shower?)

_____
_____
_____

5) *QUEM VOCÊ ACHA QUE É DONA EMA?* (Who do you think Mrs. Ema is?)

_____
_____
_____

6) *QUE NOME VOCÊ DARIA PARA ESSA HISTÓRIA?* (What name would you give to this story?)

_____
_____
_____

Scoring and instructions for individual application of the IBALEC.

1– WRITING THE ALPHABET (After writing the letters – in their conventional order or the letters the child knows – ask the student to name the letters he/she wrote from last to first. Only correctly written and named letters should be scored).

Which letters are spelled and named correctly (optional): _____
_____
_____
_____

Number of letters written and named correctly:

Scoring	Interpretation
0	Wrote and named up to 25% of the letters of the alphabet (0–6 letters)
1	Wrote and named up to 50% of the letters of the alphabet (7–13 letters)
2	Wrote and named up to 75% of the letters of the alphabet (14–19 letters)
3	Wrote and named more than 75% of the letters of the alphabet (20–26 letters)

2- WORD READING (Ask the child to read out load showing one word at a time in the order below – Note down how the word was read)

	Answer	Score	Interpretation
BIFE			0: child refuses to read or gives a random response, saying any word.
SALA			1: child identifies a syllable or first letter of the word and gives a wrong answer that is consistent with the identification, i.e. says the word or a sequence of letters that begins with the letter of the target word. (e.g.: reads "b", "bi" or "bola" for BIFE).
POMADA			
GELATINA			2: child identifies part of the word and adds, removes or changes the syllables (e.g.: read *salada* for SALA). Also do not observe de correctect grapheme-to-phoneme conversion due to lack of knowledge of contextual rules (e.g.: reads *"rossa"* for ROSA).
ROSA			
HOJE			3: child segments the word identifying its component syllables (e.g.: "ge-la-ti-na") showing difficulty with tonicity/accentuation (e.g.: for *ALFABETO* reads "al-fa-bê-to")
VELHO			
TÁXI			4: child rapidly identifies the word, suggesting automaticity.
CHUVA			
ALFABETO			
CONFUSÃO			
GALINHA			
FLORESTA			

Total score: _____   Classification:_____

3– WORD WRITING (Ask the child to write the name of the animal by the side of each picture. Make sure he/she properly recognizes the animals, saying their name when he/she doesn't).

	Score	Interpretation
COBRA – snake		0: child refuses to write the name or writes a sequence of letters or pseudo letters unrelated to the sound of the correct word.
CACHORRO – dog		1: child correctly writes a letter or syllable of the target word, e.g.: writes OIA (for *"formiga"*); PCOTL (for *"peixe"*); CAXO (for *"cachorro"*).
PÁSSARO – bird		
FORMIGA – ant		2: child writes the word legibly but with more than one spelling error, e.g.: FOMICA (for *"formiga"*); CUELO (for *"coelho"*)
PEIXE – fish		3: child writes the word with only one spelling error, e.g.: writes CACHORO (for *"cachorro"*) and/or does not accentuate the word correctly (e.g.: writes PASSÁRO, LEAO).
COELHO – rabbit		4: child writes the word without any spelling errors.
LEÃO – lion		

Total score: _____   Classification:_____

4– SENTENCE WRITING (Ask the child to write something about each picture, making sure he/she recognizes each picture).

Example: (Should be read and shown just once to make sure that the child doesn't repeat the same model for each picture).

_____   In the zoo people love to visit the giraffe.

	Syntactic structure* (0–1)	Lexical segmentation (0–1)	Spelling/Punctuation (0–0.5) / (0–0.5)		Total score (0–3)	Interpretation
PALHAÇO – clown						**Syntactic structure:** sentence word structure, nominal and verbal concordance.
CHAPÉU – hat						**Lexical segmentation:** check for hyposegmentation errors (word combination) and/or hypersegmentation.
CAMISA – shirt						
ELEFANTE – elephant						**Spelling/Score:** check adequacy of spelling and punctuation.
ESCORREGADOR – slide						

Total score: _____   Classification:_____

5– SENTENCE COMPREHENSION (The student should read the sentence and point to the corresponding picture). Each correctly understood/answered sentence receives a point.

Sentence	"O sapo pula na lagoa".	"Fui na feira com minha tia e vi muitas frutas e verduras".	"O motorista deste ônibus é o senhor João".	Total score	Interpretation
Correct answers					0: wrong answer. 1: right answer.

Total score: _____   Classification:_____

6– TEXT COMPREHENSION (The child should read the text out load but can read the text to him/ herself to answer the questions. The questions should also be read by the child. If the child is too tired to provide a written answer, let him/her answer verbally and write down the answers).

ITEM	SCORE				TOTAL
	0	**1**	**2**	**3**	
Question 1 Por que o senhor Antônio ligou a mangueira de água no jardim?	Doesn't answer	Gives an answer that is incorrect from the point of view of the information contained in the text. e.g.: *Para molhar as plantas* (to water the plants)	Considers only one of the ideas mentioned for a score of 3.	*Porque estava muito quente* (because it was really hot) (1). *Para os filhos se molharem* (So his children could get wet) (2).	
Question 2 Por que a garota entrou no quintal?	Doesn't answer	Gives an answer that is incorrect from the point of view of the information contained in the text. e.g.: *Para tomar guaraná* (To drink soda).	Considers only one of the ideas mentioned for a score of 3.	*Porque Marcos e Ana corriam e pulavam no jardim, dando gargalhada* (Because Marcos and Ana ran and jumped around the garden, laughing out loud) (1). *Porque a brincadeira das crianças chamou a atenção dos amiguinhos da rua* (Because the children's game caught the attention of their friends in the street) (2).	
Question 3 Por que as crianças decidiram tomar um banho?	Doesn't answer	Gives an answer that is incorrect from the point of view of the information contained in the text. e.g.: *Porque o pai mandou* (Because their dad told them to).	Considers only one of the ideas mentioned for a score of 3.	*Porque elas estavam cansadas de brincar* (Because they were tired of playing) (1). *Porque elas estavam sujas e molhadas* (Because they were wet and dirty) (2).	
Question 4 O que as crianças fizeram depois do banho?	Doesn't answer	Gives an answer that is incorrect from the point of view of the information contained in the text. e.g.: *Foram dormir* (They went to bed).	Considers only one of the ideas mentioned for a score of 3.	*Tomaram guaraná geladinho* (They had a cold soda) (1) *Se divertiram (foi uma festa!)* (They had fun [it was party!]) (2).	
Question 5 Quem você acha que é Dona Ema?	Doesn't answer	Gives an answer that is incorrect from the point of view of the information contained in the text. e.g.: *Responder que é algum objeto, animal* (Answers that she is an object or animal).	Give a possible answer that is out of context. e.g.: *Responder que é a professora*.	*A mãe das crianças, a esposa do Sr. Antônio, a avó, a tia* (The children's mother, Mr. Antonio's wife, grandma, aunt).	
Question 6 Que nome você daria para essa história?	Doesn't answer	Gives an answer that is incorrect from the point of view of the information contained in the text, e.g.: *Festa de Aniversário* (Birthday party).	Very long (like the formation of a sentence or summary of the story) or very short titles, e.g.: *As crianças.* (The children)	Short titles related to the overall context of the story, e.g.: *Brincadeiras no jardim; Domingo divertido; Brincando com água no jardim* (Playing in the garden; Fun Sunday; Playing with water in the garden).	

Total score: _____   Classification:_____

Contribution of each author to the work:

Gislaine Gasparin Nobile: Data collection, contributed throughout the writing process, reference search, and data analysis.

Flávio Rebustini: Co-supervisor of the research, performed all the psychometric analysis, describing the statistical part.

Sylvia Domingos Barrera: Supervisor of the research, guided the construction of the Instrument and participated in the entire writing and guidance process.