anales de psicología / annals of psychology
2025, vol. 41, nº 1 (may), 158-171
https://doi.org/10.6018/analesps.615011

Published by Editum, Servicio de Publicaciones, Universidad de Murcia (Spain), in https://revistas.um.es/analesps
ISSN online: 1695-2294. License Creative Commos 4.0 BY

Assessing the factorial validity and longitudinal invariance of indicators of basic early reading skills in Spanish first graders

Juan E. Jiménez*, and Eduardo García Miranda

The University Institute of Neuroscience (IUNE), Universidad de La Laguna, The Canary Islands, Spain.

Título: Evaluación de la validez factorial y la invarianza longitudinal de los indicadores de habilidades básicas de lectura temprana en alumnos españoles de primer grado.

Resumen: Se realizó un estudio para analizar la validez factorial y la invarianza longitudinal de medida de los indicadores de progreso de aprendizaje en lectura (IPAL) como medición basada en el currículo (MBC) en estudiantes españoles de primer grado de primaria. El modelo propuesto es un modelo de primer orden compuesto por dos factores correlacionados: un factor compuesto por las tareas de fluidez en conocimiento alfabético, fluidez en conciencia fonológica y conocimiento del lenguaje impreso, que sirven como indicadores observables para el factor subyacente de habilidades tempranas emergentes, y fluidez en lectura oral, lectura de pseudopalabras, y textos mutilados que sirven como indicadores observables para el factor de habilidades de alfabetización convencionales. El IPAL está compuesto de tres formas paralelas (i.e., A, B y C), y se administró tres veces diferentes a 947 estudiantes españoles de primer grado durante el año escolar (otoño, primavera y verano). Se concluye que la MBC muestra una validez de constructo adecuada y una equivalencia de medición, lo que permite la evaluación de habilidades esenciales de primer grado que acontecen mientras los niños aprenden a leer.

Palabras clave: Educación primaria. Lectura. Invarianza factorial longitudinal. Evaluación basada en el currículo.

Abstract: A study was conducted to analyze the factor validity and longitudinal measurement invariance of indicators of basic early reading skills in Spanish first-grade students via a curriculum-based measurement (CBM) approach. The proposed model is a first-order model comprising two correlated factors: one factor, comprising tasks measuring alphabetic knowledge, phonological awareness, and print awareness, serves as an observable indicator for the foundational skills underlying factor, whereas oral reading fluency, pseudoword reading, and maze sentences serve as observable indicators of the conventional literacy or reading skills factor. The CBM consisted of three parallel forms (A, B, and C) and was administered at three different times during the school year (fall, winter, spring) to 947 Spanish first-grade students. It is concluded that the CBM demonstrates adequate construct validity and measurement equivalence, allowing for the evaluation of first-grade essential skills that occur as children learn to read.

Keywords: Elementary grade. Reading. Longitudinal measurement invariance. Curriculum-based measurement.

Introduction

One of the greatest challenges for schools is ensuring that students learn to read and develop the ability to read to acquire knowledge. When students fail to develop competent reading skills by the end of the third grade, they encounter various challenges that adversely affect their academic performance. By the fourth grade, students are confronted with diverse texts, and those who are more skilled readers have acquired the ability to extract and analyze new information, enabling them to enrich their vocabulary through reading. However, students who do not achieve this goal are more likely to drop out of school (O'Brien, 2008). In Spain, Ley Orgánica 3/2020 (LOMLOE) established early detection and intervention of reading difficulties as a primary objective in primary education, emphasizing the development of phonological awareness, letter recognition, decoding skills, reading fluency, and comprehension as key literacy competencies (Ministerio de Educación, 2020). These foundational and conventional literacy skills are essential for ensuring that students acquire proficient reading abilities during the early years of primary education. One of the most effective approaches for the early identification of reading difficulties and progress monitoring is curriculum-based measurement (CBM), a formative assessment method designed to evaluate key literacy skills in alignment with educational curricula (Jenkins et al., 2013). CBMs are designed for continuous progress monitoring, providing educators with frequent, reliable data to support instructional decision-making and early intervention. Unlike standardized assessments such as Prolec-R (Cuetos et al., 2007) and BIL (Sellés et al., 2010), which are designed for a one-time diagnostic evaluation of reading skills or reading precursors, CBM functions as a formative assessment tool for ongoing progress monitoring within the school context. While standardized tests provide a static profile of a student's reading ability, CBM facilitates frequent evaluations, delivering immediate and actionable insights for teachers to tailor instruction and intervention strategies.

The local curriculum of the Canary Islands defines the condition of "specific learning difficulties in reading or dyslexia" as "poor performance in the lexical processes involved in reading... characterized by difficulties in decoding isolated words that generally reflect insufficient phonological processing skills... the lack of automation of lexical processes hinders adequate fluency, which may contribute to poor reading comprehension." This definition largely coincides with that proposed by the International Dyslexia Association (IDA) (2002). However, neither Spanish nor local legislation defines specific strategies for identifying children at risk of reading disabilities (RDs) early on, and students can only be identified with RDs after two years of failure. In recent years,

*** Correspondence address [Dirección para correspondencia]:**
Juan E. Jiménez. E-mail: ejimenez@ull.edu.es
(Article received: 09-05-2024; revised: 09-01-2025; accepted: 19-02-2025)

this wait-to-fail model has been replaced by a response to intervention (RtI) model (Jiménez, 2019).

The process of learning to read is not simple, as it involves multiple components in the development of reading ability. A prominent theory in the realm of literacy acquisition is Frith's model of reading development (Frith, 1985). According to this model, children progress through distinct stages, each associated with age and experience. Her model delineates three stages: the logographic stage, where children recognize symbols or logos instantly; the alphabetic stage, where they start associating letter symbols with sounds in words, often with partial accuracy; and the orthographic stage, where they internalize spelling patterns and read words more automatically. Other theoretical models of reading, such as the simple view of reading (Gough et al., 1996; Gough & Tunmer, 1986), the dual route cascading model (Colheart et al., 2001), and the triangle model of reading (Harm & Seidenberg, 1999, 2004; Plaut et al., 1996), provide a conceptual framework suggesting that decoding processes are fundamental in the early stages of reading development. Precise and fluent word decoding is one of the necessary components or prerequisites for achieving the main goal of reading: obtaining the overall meaning of the text.

Main Precursors and Components of Reading

Two prominent expert reports, the National Reading Panel (NRP) (2000) and the National Early Literacy Panel (NELP) (2008), detail the crucial abilities needed for reading. They distinguish between conventional literacy skills (decoding, fluency, comprehension, writing, spelling) and foundational skills, which are precursor abilities indicating the development of conventional literacy skills. These early literacy skills, including alphabet knowledge, phonological awareness, and rapid naming, consistently predict later literacy development. Additionally, skills such as concepts about print and oral language also contribute to early literacy.

Indeed, an early literacy skill identified by the NELP is print awareness. Clay (1985) suggested that children who struggle to grasp the concepts of print and the instructional language used by teachers while reading lessons might experience confusion and difficulties at an early stage in their journey of learning to read and write. This foundational literacy skill encompasses understanding the basic features and functions of written language. Children's development of print awareness involves recognizing that print carries meaning, understanding the distinction between letters and words, comprehending the directionality of text (e.g., left to right, top to bottom), and recognizing the spaces between words (Papadopoulos & Bourogianni, 2024). Print awareness is considered an early literacy skill that is usually more predictive of literacy achievement at the end of kindergarten or the beginning of first grade than of later literacy growth (NELP, 2008).

Both the NRP and the NELP recognized phonemic awareness (PA) as one of the five essential components of reading and concluded that explicit instruction is effective in fostering PA skills in students. Phonological awareness refers to a metalinguistic skill enabling the detection, manipulation, or analysis of auditory aspects of spoken language without considering their meaning (Mariángel & Jiménez, 2016). Notably, oral language can be segmented into various units (e.g., syllables, onset, rhyme, and phonemes), allowing us to establish connections between oral and written language. Phonemic awareness has been shown to be the most predictive level of reading ability in Spanish (Cuadro & Trías, 2008; González-Valenzuela et al., 2016; Serrano et al., 2005). Moreover, on the basis of three recent meta-analyses (Baker et al., 2022; Rehfeld et al., 2022; Rice et al., 2022), it is evident that phonemic awareness (PA) instruction plays a crucial role in enhancing reading ability across different populations and languages.

Another component considered in both NRP and NELP is alphabet knowledge, which encompasses children's ability to identify the names and sounds of alphabet letters (Puranik et al., 2011). It has been demonstrated to be a prominent predictor of reading acquisition, both in the short and long term, across various orthographic systems (Caravolas et al., 2013; Lervåg et al., 2009; Torppa et al., 2016). Particularly relevant in transparent alphabetic systems, where grapheme–phoneme correspondence rules are highly regular, such as in Spanish, this knowledge facilitates reading through the phonological or indirect route mentioned in previous theoretical models. Research has shown that rapid acquisition of letter knowledge contributes to reading fluency, even after phonological skills are controlled for (Torppa et al., 2016). As the correspondence between letters and sounds is unambiguous, children can quickly identify and encode letters within words (Acha et al., 2023). Furthermore, recent studies have highlighted the importance of grapheme learning sequences in Spanish, emphasizing that they differ from phoneme learning sequences, unlike in English (de la Calle et al., 2018). This study underscores that a specific sequence in grapheme learning is crucial for predicting decoding skill development. This study concludes that early educational practices considering both the name and sound of graphemes, along with a determined sequence in their learning, enhance reading performance in children learning to read in Spanish.

The vocabulary component is considered in the NRP and refers to the set of words necessary for effective communication. It is logical that vocabulary plays a critical role in reading comprehension, as encountering many unfamiliar words in a text affects comprehension (Kim & Pallante, 2012). The simple view of reading (SVR) proposes that essential knowledge for reading comprehension lies in the mastery of vocabulary (Hoover & Gough, 1990). This vocabulary knowledge, which is seen as the basic unit in reading comprehension and as a key element in language understanding, involves the ability to form mental images of passages and to identify the semantic meaning of individual words or characters (Braze et al., 2016; Nation, 2015). According to a reciprocal model, understanding vocabulary facilitates text comprehension, and in turn, text comprehension contributes to vocabulary learning (Stanovich, 1986). Students with extensive vocabularies often comprehend texts better, as they are likely to understand the meanings of the words contained in them. In contrast, those with limited vocabularies often struggle to comprehend the text and tend to read less, resulting in lower incidental learning of new words.

The fluency component in reading is also considered in the NRP and refers to the ability to read accurately, appropriately, and with expression or prosody (Khun & Stahl, 2003). Recent studies have shown that children read with expression only when they have achieved a high level of reading automatization (Benjamin & Schwanenflugel, 2010; Rodriguez et al., 2024; Schwanenflugel et al., 2004; Veenendaal et al., 2016). In this context, studies in Spain highlight that a pattern of prosody similar to that of adults is observed from the fifth grade onward (Álvarez-Cañizo et al., 2018). Reading this way allows us to focus our attention on text comprehension (Khun & Stahl, 2003; NRP, 2000). Automatization in word recognition is an essential component of expert reading, resulting in speed, fluency, and cognitive resource release (Schwanenflugel et al., 2004). This ability is particularly critical in transparent languages such as Spanish, both in assessment and instruction, owing to the simplicity of learning phoneme–grapheme rules. However, some studies suggest that in languages with transparent orthographies, the strong relationship between decoding and prosody may weaken over time, unlike what happens in opaque orthographies. For example, Calet et al. (2015) reported differences in the relationship between word reading automatization and prosody in Spanish second- and fourth-grade students, with the magnitude of this relationship being lower in fourth-grade students. This suggests that as readers become more proficient in decoding, other skills, such as syntactic skills, may also influence prosody development (Rodríguez et al., 2024). In fact, to read a sentence with intonation, it is necessary to assign syntactic roles to words in a sentence; therefore, Kuhn and Stahl (2003) proposed that reading with proper expression requires not only solid decoding skills but also an understanding of the syntactic structures of the text. In this context, Rodríguez et al. (2024) studied word reading skills and syntactic knowledge in fifth-grade Spanish-speaking students. They reported that word reading automatization strongly influenced reading prosody, whereas syntactic knowledge had a weaker effect, possibly due to the lack of sensitive measures in early primary grades.

Finally, the reading comprehension component considered in the NRP is defined as the ability to interpret text information and, more broadly, as the skill that allows students to access specific domains of knowledge and provide a means to achieve both intellectual and affective goals (Snow, 2002). Recently, Baker et al. (2022) conducted a meta-analysis aiming to synthesize studies investigating the relationship between reading components prescribed by the NRP and reading comprehension. The main findings highlight the importance of phonemic awareness and the alphabetic principle in reading comprehension in Spanish. Additionally, the findings emphasize that the relationship between reading comprehension and word reading in isolation has the greatest impact. Moreover, the transparency of orthography does not guarantee automatic word reading, as vocabulary also influences this skill and reading comprehension.

In summary, these findings emphasize the need to consider multiple key components when designing curriculum-based assessments to detect and support students at risk of learning difficulties in reading.

Curriculum-based measurement

CBM has emerged as a valuable tool for the early identification of students at risk for reading difficulties and the monitoring of literacy development in young learners (Catts et al., 2015; Christ et al., 2013). In this context, IPAL was developed as a CBM specifically tailored to the Spanish educational curriculum, ensuring its alignment with curricular objectives and incorporating input from early literacy experts to align with instructional goals. As a result, IPAL qualifies as a CBM by directly assessing key literacy skills and providing educators with reliable data to track student progress over time. Moreover, the design of IPAL adheres to the core principles of CBMs by incorporating short, fluency-based tasks that can be administered multiple times a year (fall, winter, and spring) to monitor students' progress. This structure enables continuous progress monitoring, capturing students' reading development across different time points. Each form includes fluency-based tasks that assess both foundational and conventional literacy skills, such as alphabetic knowledge, phonological awareness, decoding, and oral reading fluency. Unlike standardized reading tests that provide a one-time assessment, this repeated measurement approach differentiates IPAL from traditional standardized tests, making it a practical tool for teachers seeking to identify struggling readers early and implement evidence-based interventions within an RtI framework.

CBM serves as a formative assessment method with the dual purpose of early detection of students at risk of learning difficulties (LDs) and monitoring learning progress (Catts et al., 2015; Christ et al., 2013). According to these authors, this type of assessment consists of a set of measures that utilize the general education curriculum for test development. The use of CBM is increasingly recommended for universal screening from early childhood education to the second or third grade of primary education. CBMs should be brief and quickly administered, with each measurement lasting between 1 and 5 minutes, depending on the specific skill being assessed. Providing students with evidence-based, high-quality instruction is best achieved through highly personalized assessment methods tailored to each individual's learning process rather than relying primarily on standardized test batteries that often inaccurately reflect students' actual competence levels (Jiménez, 2019). In this context, the IPAL was designed as a CBM specifically tailored to the Spanish educational system. It consists of three parallel forms (A, B, and C) administered at three different time points during the academic year (fall, winter, and spring). This structure allows for continuous progress monitoring, capturing students' reading development over time. Each form includes fluency-based tasks that assess foundational and conventional literacy skills, such as alphabetic knowledge, phonological awareness, decoding, and oral reading fluency. The repeated measurement design of IPAL aligns with the CBM framework, providing educators with frequent, reliable data to guide instruction and early intervention efforts. This feature makes it a practical and efficient tool for identifying at-risk students and tracking their response to instruction within an RtI model.

Standardized tests, as outlined by Stecker et al. (2008), may compare students to a norm and provide a clear view of their performance levels compared with peers, but their administration requires a considerable amount of time. Additionally, these tests seldom capture classroom-taught content and offer limited information for teachers to make instructional decisions. In contrast to traditional assessments, which utilize standardized tests, the most common instruments used in prevention approaches, such as the RtI model, are universal screening and progress monitoring (Jiménez, 2019). Universal screening, established as the critical first step in this preventive approach, identifies students who may be at risk of future academic difficulties in the evaluated academic area and thus would benefit from more personalized and tailored instruction. These screening assessment tools must accurately identify these students while demonstrating their technical adequacy and being cost-effective and easy to administer (Glover & Albers, 2007). Furthermore, progress monitoring, a key component of CBM, integrates assessment and instruction to enhance student learning. CBM represents a modern evaluation concept that unifies these elements rather than introducing a new one. These procedures originated in the Anglo-Saxon context amid professional and political changes (Shinn et al., 1989), as the need to better equip teachers with functional tools to measure academic progress became increasingly evident. Among the problems faced by teachers were the lack of connection between assessment and classroom instructional materials, as well as the insensitivity of tests used to capture student learning growth.

Purpose of the Study

To our knowledge, validated CBM tools for assessing the main precursors and components of reading in first-grade Spanish-speaking students are lacking. While some CBM tools in Spanish, such as Indicadores Dinámicos del Éxito en la Lectura (IDEL; Baker et al., 2007), have been developed to assess early literacy skills, research examining their construct validity and longitudinal factorial invariance remains scarce. Although CBM is a well-established assessment approach in English, its psychometric properties, particularly its factorial structure and measurement stability over time, have been less frequently examined in Spanish. In a systematic review of CBM research in languages other than English, Vázquez et al. (2024) reported that despite the growing use of CBMs in Spanish-speaking populations, studies focusing on their psychometric properties—such as construct validity, factorial structure, and measurement invariance—are still limited. This gap in the literature highlights the need for studies such as the present one, which seeks to provide empirical evidence of the construct validity and longitudinal stability of CBM IPAL in Spanish-speaking first-grade students. The lack of research on whether CBMs maintain measurement consistency over time limits their potential to accurately track literacy development, reducing their effectiveness in guiding instructional decisions. By facilitating continuous progress monitoring and incorporating a rigorous validation framework, CBMs such as IPAL ensure that students' literacy development remains aligned with instructional expectations, enabling timely and effective educational decision-making. Gutiérrez et al. (2022) analyzed the technical adequacy, diagnostic accuracy, and best predictors of several curriculum-based assessment measures that are part of IPAL. Growth slopes were analyzed through mixed models, and the predictive validity of this CBM was examined through binary logistic regression models. Most of the measures included in the IPAL study exhibited adequate reliability and were sensitive to change, as differences in growth slopes were found according to risk status. However, construct validity and its stability over time have not been studied thus far. The latter is important when the CBA measure is used to monitor learning progress.

Therefore, the aim of this study is to assess a hypothetical model aimed at gaining a deeper understanding of the multidimensional nature of reading proficiency among first-grade elementary school students. We have incorporated skills traditionally addressed in the literature, including both foundational skills and conventional literacy skills. Given our focus on first grade, we anticipate that the skills encompassed by the IPAL instrument will cluster around a first-order factor. To achieve this goal, we employ both exploratory factor analysis (EFA) to explore the underlying structure of the data and confirmatory factor analysis (CFA) to confirm the proposed model. Specifically, CFA will be utilized to determine the extent to which the variance associated with each task comprising the IPAL instrument is explained by a single underlying factor. Additionally, factorial invariance analysis examines the degree of measurement stability over time. We anticipate that one underlying factor will emerge, encompassing both foundational skills and conventional literacy skills. This first-order factor is expected to be theoretically grounded and reflective of the developmental dimensions typical for first-grade elementary students. This analysis provides insights into the consistency of the identified factor structure across different time points, enhancing our understanding of reading proficiency development in early elementary education.

Method

Participants

A comprehensive study advocated by the Department of Education of the Canary Islands Government, in collaboration with Universidad de La Laguna, as a large-scale project across all the islands for piloting the RtI model involved the participation of 77 public schools situated in the Canary Islands (Jiménez et al., 2021). Among these, 62 schools situated across the seven islands of the Autonomous Community of the Canary Islands were assigned to the experimental group, whereas 15 schools located on the capital islands, namely, Tenerife and Gran Canaria, composed the control group. To ensure comparability between both groups in terms of geographical distribution (urban, suburban, or rural) and socioeconomic status, the selection and assignment of schools were carried out by the Canarian Agency for University Quality and Educational Evaluation (ACCUEE). While the experimental group included schools from all seven islands, the control group was composed of schools from the capital islands, Tenerife and Gran Canaria. However, while the selection of geographical and socioeconomic strata was structured, the allocation of schools to the experimental and control groups within each stratum was conducted randomly. The sample for the current study comprised 947 students enrolled in the 1st grade of primary education who were part of the control group. The mean age of the participating students at the start of the test was 75.55 months (6.3 years), with a standard deviation of 3.99.

Instruments

The materials used were based on the IPAL (Jiménez & Gutiérrez, 2019). There were three versions (Form A, Form B, Form C) that were equivalent in difficulty level but with different items, each tailored to the school curriculum. The tasks included in the IPAL assess fundamental reading skills taught in first grade, such as alphabet knowledge, phonological awareness, decoding, and oral reading fluency. To ensure its content validity, IPAL was explicitly designed to align with the literacy competencies outlined in Ley Orgánica 3/2020 (LOMLOE). The LOMLOE emphasizes the progressive acquisition of reading skills, ensuring that students develop phonological awareness, letter recognition, decoding skills, and reading fluency—key areas that the IPAL tasks assess. The letter name and sound knowledge task measures students' ability to recognize letters and associate them with corresponding sounds, a fundamental literacy skill emphasized in the curriculum. The phonemic awareness task evaluates students' ability to segment spoken words into phonemes, which is essential for decoding and spelling development. The print awareness task assesses students' understanding of basic text conventions, such as directionality and punctuation, which are critical components of early reading instruction. The pseudoword reading task measures grapheme-phoneme correspondence skills, reinforcing phonological decoding strategies that are a core part of literacy instruction. The cloze text task evaluates sentence-level reading comprehension, assessing students' ability to infer meaning from context and supporting vocabulary development. Finally, the oral reading fluency task assesses reading accuracy and speed, reflecting automaticity in word recognition and fluency development. This explicit alignment with the first-grade reading curriculum ensures that IPAL measures the core competencies that students are expected to acquire and develop over the academic year. These measures are as follows:

Letter name knowledge (FNL): This task requires children to name as many uppercase and lowercase letters as possible in one minute. It included twenty-seven letters of the alphabet randomly ordered on a sheet of paper with one hundred letters. The raw score was the number of letter names identified correctly within the minute (ICCr = .99, p < .001).

Letter sound knowledge (FSL): This task was presented on the same sheet as FNL but required the students to identify the correspondence between letters and sounds as quickly as possible in one minute. The raw score was the number of letter sounds identified correctly within the minute (ICCr = .99, p < .001).

Phonemic Awareness (PA): This task requires students to segment nonsense words presented orally into their phonemes as quickly as possible in one minute. It consists of twenty nonsense words instead of real words to avoid segmentation through the orthographic route. There are four levels of difficulty based on the syllabic structure: CVCV (e.g., tapu), VCV (e.g., ura), CVCVC (e.g., patro), and CVCCV (e.g., clepa). The stimulus presentation was divided into five blocks presented sequentially. Each block contained one word from each difficulty level, and they were ordered by difficulty level. The raw score was the number of phonemes identified correctly within the minute (ICCr = .99, p < .001).

Print awareness (PRA): This task evaluated basic knowledge about what printed language represented in general through six questions that the student had to answer while observing a textbook suitable for first grade. The questions included aspects related to book sections (e.g., [1] "Point where the book starts") and directionality (e.g., [2] "Where should I continue reading after this sentence? "), terms related to print (e.g., word, letter, beginning and end of a sentence; e.g., [3] "Point to a capital letter"; [4] "Point to the beginning and end of a sentence"; [5] "Point to where there are two different sentences"), and punctuation and exclamation marks (e.g., pointing to a period, ask: [6] "What is the name of this symbol? What is it used for?"). The raw score was the number of aspects answered correctly (a maximum of six). The raw score was the number of questions answered correctly (ICCr = .98, p < .001).

Pseudowords Reading (PSR): This task requires students to read as many pseudowords as possible in one minute. Its objective was to measure two basic components of word reading: grapheme-phoneme correspondence rules and the ability to combine letter sounds when reading unfamiliar words. Forty pseudowords were used to assess the phonological route. The raw score was the number of pseudowords read correctly within the minute (ICCr = .98, p < .001).

Cloze Text (OCT): This subtest was a five-minute evaluation of vocabulary and comprehension at the phrase level. The student read twenty incomplete phrases where the last word had been removed and replaced by three multiple-choice alternatives. They had to then select the one that best completed the sentence (e.g., "The slowest animal is the: Snail - Lion - Tiger"). The raw score was the number of correct answers identified within five minutes (ICCr = .99, p < .001).

Oral reading fluency (ORF): This is a one-minute fluency assessment in which students are asked to read aloud a text with speed and accuracy. It consists of a short story of one hundred thirty-three words composed of three paragraphs that progressively increase in difficulty depending on the type of connector and the syllabic structure of the words. The raw score was the number of words read correctly in one minute (ICCr = .96, p < .001).

A composite score was derived from measures of letter sound knowledge and letter name knowledge to assess alphabetic knowledge.

Procedure

This study was conducted in accordance with the guidelines of the Ethics Committee of the University of La Laguna (https://viinv.ull.es/ceiba/). Prior to administering the instrument, authorization was obtained from the school and families for the study. In each of the capital islands, seven psychologists and two educational psychologists were trained in the administration of the IPAL tool to collect the information. The CBMs were administered at three different time points throughout the school year (fall, winter, and spring), with an approximate interval of three months between each administration. The assessments were conducted individually during school hours.

Analysis of Data

To avoid overfitting and enhance generalizability, the fall dataset was randomly split into two independent subsamples (n_1 = 474; n_2 = 473), with one used for exploratory factor analysis (EFA) and the other for confirmatory factor analysis (CFA). This approach allows for the identification of the underlying factor structure in the first half of the data and its subsequent confirmation in the second half. The feasibility of factorizing the data matrix was assessed via both the Kaiser–Meyer–Olkin (KMO) criterion and Bartlett's test of sphericity. Parallel analysis was then employed to determine the appropriate number of factors to retain. An exploratory factor analysis (EFA) was conducted on the first half of the fall dataset to discern the underlying factors within the CBM. Given that the data met the assumptions of linearity and that outliers were controlled, the maximum likelihood method was selected as the extraction method, as it allows for the calculation of a wider range of fit indices and remains robust when the data are not normally distributed. The number of factors to extract was determined via parallel analysis and examination of the scree plot. Given the theoretical expectation of correlated factors and the need to ensure parsimony and theoretical coherence, an oblimin rotation method was adopted (Child, 2006). To assess the stability of the factor structure and item assignments, a bootstrap analysis with 500 iterations was performed. In each iteration, a bootstrap sample (sampling with replacement) was drawn from the EFA dataset, and an EFA with the same parameters as described above was conducted. The frequency with which each item was assigned to each factor across the 500 iterations was calculated, and an item stability index (proportion of times assigned to the most frequent factor) was computed. The measurement model was subsequently evaluated via confirmatory factor analysis, which was executed separately for each time point (fall, winter, and spring). In each instance of the confirmatory factor analysis, the variances of the error terms were designated as free parameters, whereas a structural coefficient in the latent variable or factor was set to one. The maximum likelihood robust (MLR) method was used for estimation. Reliability and convergence were assessed through calculations of Cronbach's alpha coefficient (α), hierarchical omega (ωh), and average variance extracted (AVE). Model fit was determined in accordance with recommendations from Hu and Bentler (1999), employing absolute fit measures, including the chi-square statistic and root mean square error of approximation (RMSEA), and incremental fit indices, such as the Tucker–Lewis index (TLI), standardized root mean square residual (SRMR), and comparative fit index (CFI). Finally, the longitudinal invariance of the IPAL variable across the three measurement points was estimated using the entire sample. This involved a progressive approach based on various levels of restriction (Milfont & Fischer, 2010) using standardized scores from the IPAL tasks. The steps included confirming the equivalence of basic measurement models for each time point (configural model), verifying the invariance of factor loadings over time (metric model), enforcing uniformity in intercepts across time (scalar model), and enhancing scale comparability by imposing restrictions on the error variance of the residuals of the indicators (strict model). The criteria for evaluating reliability, validity, and model fit remained consistent with those of the previous stages. Evaluation of longitudinal factorial invariance was conducted through mean comparison in nested models (Newsom, 2015) and assessment of incremental fit indices (Milfont & Fischer, 2010). All analyses were conducted via the R program (R Core Team, 2022) along with the lavaan (Rossell, 2012), semTools (Jorgensen et al., 2022), and psych (Revelle, 2023) packages. To ensure the representativeness of the data, multivariate extreme cases (Mahalanobis distance) and subjects with any missing values were excluded.

Results

Table 1 presents the descriptive statistics of the scores for each IPAL task, including the maximum and minimum values, means, standard deviations, skewness, and kurtosis, for each task across the measurement points. The skewness and kurtosis values of all the study variables indicate that there are no issues with the distribution of the data, as the absolute values of these indices do not exceed 3.0 and 10.0, respectively (Kline, 2005). The descriptive statistics suggest greater stability in print awareness across measurement points, unlike the other components, which show more variation over time.

The correlation matrix of the results obtained in the different IPAL tasks according to the measurement point is presented in Table 2. The results demonstrate statistically significant correlations among all the indicators.

Exploratory Factor Analysis

EFA feasibility was verified by assessing the factorization viability of the data matrix via Kaiser–Meyer–Olkin (KMO = .84) and Bartlett's test of sphericity ($\chi2(15) = 2153.11$; $p < .001$). A parallel analysis was conducted to determine the optimal number of factors to extract in the factor analysis, and 500 random datasets were generated for the comparison of eigenvalues. Parallel analysis suggested that two factors should be retained. The eigenvalues of the first two factors (3.36 and .25) were greater than the corresponding average eigenvalues from the simulated datasets (.54 and .09). The remaining factors had eigenvalues very close to zero or negative, smaller than or similar to the simulated values. Based on these, an EFA with a two-factor solution was conducted. EFA revealed the existence of these factors, which explained 63% of the total variance. The factors were labeled as foundational skills (FS) and conventional literacy skills (henceforth, reading skills) (RS). Factor FS exhibited high loadings in tasks related to alphabet knowledge (AK), phonological awareness (PA), and print awareness (PRA), whereas Factor RS showed such loadings in tasks involving pseudoword reading (PR), cloze text (CT), and oral reading fluency (ORF). Table 3 presents the estimated model via EFA, including factorial loadings, item communalities, uniqueness, complexity, and Kaiser–Meyer–Olkin sampling adequacy measures for each task of the IPAL.

Table 1
Descriptive statistics of IPAL task performance across measurement time points

	Mean	*SD*	Min	Max	Skewness	Kurtosis
			Fall			
AK	50.78	25.41	0	139	0.43	0.11
PA	17.67	14.55	0	85	1.48	2.62
PRA	4.87	1.19	0	6	-1.15	1.27
PSR	11.90	9.33	0	40	0.51	-0.67
CT	7.32	5.91	0	20	0.28	-1.26
ORF	28.40	19.97	0	114	0.79	0.51
			Winter			
AK	71.93	27.27	0	152	-0.13	0.00
PA	23.03	13.20	0	80	0.78	1.00
PRA	4.37	1.22	0	7	-0.95	1.23
PSR	20.44	9.49	0	40	-0.33	-0.38
CT	9.25	4.93	0	20	-0.30	-0.75
ORF	41.44	19.71	0	113	0.02	0.38
			Spring			
AK	83.86	27.42	0	168	-0.23	0.34
PA	29.35	14.60	0	85	0.62	0.87
PRA	4.75	1.11	0	6	-0.67	0.06
PSR	24.76	8.96	0	40	-0.72	0.50
CL	14.43	4.90	0	20	-1.45	1.46
ORF	51.86	21.93	0	133	0.14	0.59

Note: The following abbreviations refer to AK: alphabetic knowledge, PA: phonological awareness, PRA: print awareness, PSR: pseudoword reading, CT: cloze text, ORF: oral reading fluency.

Table 2
Correlations among the blinded tasks at each measurement time point

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. AK1																	
2. AK2	0.79***																
3. AK3	0.72***	0.86***															
4. PA1	0.47***	0.46***	0.42***														
5. PA2	0.58***	0.66***	0.61***	0.60***													
6. PA3	0.53***	0.62***	0.63***	0.54***	0.84***												
7. PRA1	0.41***	0.40***	0.40***	0.27***	0.26***	0.29***											
8. PRA2	0.43***	0.46***	0.45***	0.28***	0.33***	0.36***	0.44***										
9. PRA3	0.38***	0.40***	0.42***	0.25***	0.27***	0.27***	0.41***	0.52***									
10.PSR1	0.73***	0.63***	0.56***	0.46***	0.46***	0.42***	0.34***	0.36***	0.31***								
11.PSR2	0.63***	0.69***	0.63***	0.39***	0.47***	0.45***	0.37***	0.39***	0.35***	0.74***							
12.PSR3	0.53***	0.61***	0.64***	0.32***	0.40***	0.43***	0.33***	0.37***	0.34***	0.59***	0.82***						
13. CT1	0.66***	0.55***	0.51***	0.46***	0.42***	0.40***	0.35***	0.37***	0.32***	0.83***	0.67***	0.53***					
14. CT2	0.57***	0.61***	0.57***	0.37***	0.45***	0.44***	0.39***	0.44***	0.37***	0.66***	0.76***	0.66***	0.70***				
15. CT3	0.52***	0.60***	0.59***	0.30***	0.42***	0.44***	0.40***	0.41***	0.36***	0.51***	0.72***	0.76***	0.53***	0.74***			
16.ORF1	0.72***	0.61***	0.56***	0.45***	0.45***	0.40***	0.35***	0.36***	0.29***	0.89***	0.75***	0.61***	0.81***	0.68***	0.54***		
17.ORF2	0.66***	0.67***	0.62***	0.39***	0.47***	0.44***	0.37***	0.40***	0.34***	0.75***	0.86***	0.78***	0.71***	0.77***	0.71***	0.84***	
18.ORF3	0.62***	0.62***	0.62***	0.35***	0.41***	0.43***	0.35***	0.39***	0.34***	0.68***	0.79***	0.81***	0.65***	0.72***	0.71***	0.77***	0.88***

Note: The following abbrevations refer to AK: alphabetic knowledge, PA: phonological awareness, PRA: print awareness, PSR: pseudoword reading, CT: cloze text, ORF: oral reading fluency. And number following it (1,2, and 3) refers to the measurement time points: fall, winter and spring respectively.

Table 3
Exploratory factor analysis results.

Task	Factor FS	Factor RS	Communalities	Uniqueness	Complexity	KMO
AK	0.843	0.048	0.775	0.225	1.007	0.88
PA	0.619	-0.041	0.346	0.654	1.009	0.89
PRA	0.427	0.003	0.184	0.816	1.000	0.90
ORF	-0.091	0.975	0.823	0.177	1.017	0.82
PSR	0.055	0.914	0.915	0.085	1.007	0.78
CT	0.189	0.704	0.738	0.262	1.144	0.90

Note: The following abbreviations refer to FS: foundational skills, RS: reading skills, AK: alphabetic knowledge, PA: phonological awareness, PRA: print awareness, PSR: pseudoword reading, CT: cloze text, ORF: oral reading fluency, KMO: sampling adequacy measure.

The percentages of explained variance and the correlations between factors can be seen in Table 4.

The goodness-of-fit statistics for the two-factor model were acceptable. The root mean square of the residuals (RMSR) was .01, and the Tucker–Lewis index (TLI) was .993. The RMSEA was .041, with a 90% confidence interval of .000 - .083. The Bayesian Information Criterion (BIC) was -17.63. These results suggest that the two-factor model provides a good fit to the data. The bootstrap analysis demonstrated high stability in the item assignments across the 500 iterations. The item stability index, which represents the proportion of times each item was assigned to its most frequent factor, ranged from .994 to 1.000, indicating very high stability.

Table 4
Descriptive statistics, factor correlations, and explained variance from the rotated exploratory factor analysis solution.

Properties	Factor FS	Factor RS
Factorial loadings sum	1.409	2.372
Variance proportion	0.235	0.395
Cumulative Variance	0.630	0.395
Explained proportion	0.373	0.627
Cumulative	1.000	0.627
Factor FS	1.000	0.773
Factor RS	0.773	1.000

Note: The following abbreviations refer to FS: foundational skills and RS: reading skills.

Confirmatory factor analyses

The proposed measurement model for the IPAL, at each measurement point, was established on the basis of the previous analysis. The (robust) indices for each model are listed in Table 5.

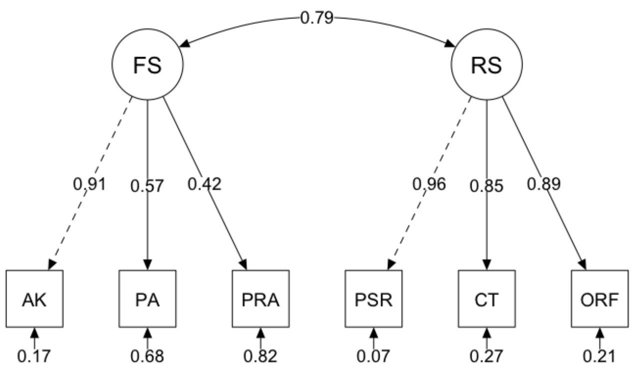
Table 5
Robust fit indices of the confirmatory factor models for each measurement time point.

Time point	χ2	df	p	nfi	nnfi	cfi	tli	rmsea	rmsea ic 90%
Fall	26.100	8	0.001	0.983	0.987	0.993	0.987	0.057	0.025-0.090
Winter	20.147	8	0.010	0.983	0.985	0.992	0.985	0.062	0.029-0.097
Spring	16.672	8	0.034	0.987	0.993	0.996	0.993	0.041	0.011-0.068

IPAL Form A. The proposed model for fall exhibits an adequate fit. The analysis results yielded the following (robust) indices for the model: $\chi 2(8) = 26.1$, $p \leq .01$, CFI = .993, TLI = .987, RMSEA = .057. The reliability of the CFA model was assessed via Cronbach's alpha, McDonald's omega hierarchical (ωh), and average variance extracted (AVE). For factor FS, the alpha coefficient was $\alpha = .574$, the omega coefficient was ωh = .803, and the AVE was .711. Factor RS showed higher reliability, with an alpha of $\alpha = .788$, ωh = .909, and an AVE of .810. The overall model demonstrated good reliability, with $\alpha = .771$ and ωh = .919, and the AVE for the total model was .746. A graphical representation of the model is depicted in Figure 1.

IPAL Form B. A goodness-of-fit assessment revealed that the proposed model for the second measurement in winter exhibited an acceptable fit. The robustness indices for the model were as follows: $\chi 2(5) = 20.147$, $p \leq .01$, CFI = .992, TLI = .985, RMSEA = .062. The reliability indices included, for Factor FS, $\alpha = .700$, ωh= .734, and AVE = .654; for Factor RS, $\alpha = .927$, ωh= .929, and AVE = .812; and for the overall model, $\alpha = .881$, ωh= .899, and AVE = .654. A graphical representation of the model is depicted in Figure 2.

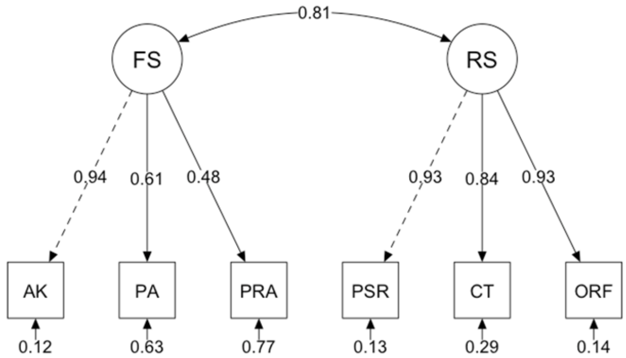
Figure 1
Factor analysis model for the measurement of IPAL form A (fall).



Note. The following abbreviations are used: FS: foundational skills, RS: reading skills, AK: alphabetic knowledge, PA: phonological awareness, PRA: print awareness, PSR: pseudoword reading, CT: cloze text, ORF: oral reading fluency.

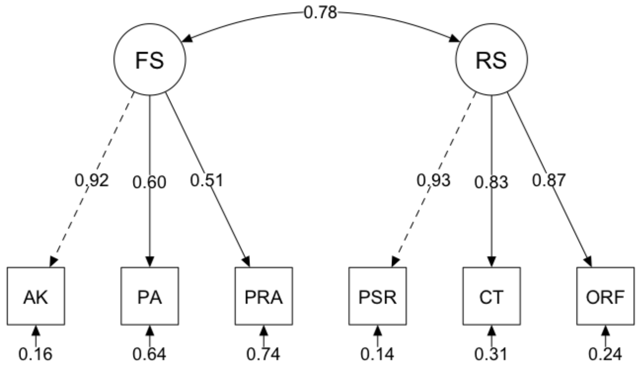
IPAL Form C. Maintaining the same measurement model as in the previous two instances, we also observed that the model aligned well with the data, as indicated by $\chi 2(8) = 16.672$, $p \leq . 5$, CFI = .996, TLI = .993, RMSEA = .041. The reliability indices for Factor FS were α = .659, ωh = .712, and the AVE was .501. For Factor RS, α= .918, ωh = .922, and the AVE is .822. The overall model included α = .864, ωh = .892, and AVE = .639. Consequently, following the assessment of goodness-of-fit, we concluded that the confirmatory factor analysis model for the winter measurement was appropriate. A graphical depiction of the model is presented in Figure 3.

Figure 2
Factor analysis model for the measurement of IPAL form B (winter).



Note. The following abbreviations are used: FS: foundational skills, RS: reading skills, AK: alphabetic knowledge, PA: phonological awareness, PRA: print awareness, PSR: pseudoword reading, CT: cloze text, ORF: oral reading fluency.

Figure 3
Factor analysis model for the measurement of IPAL form C (winter).



Note. The following abbreviations are used: FS: foundational skills, RS: reading skills, AK: alphabetic knowledge, PA: phonological awareness, PRA: print awareness, PSR: pseudoword reading, CT: cloze text, ORF: oral reading fluency.

Longitudinal factorial invariance

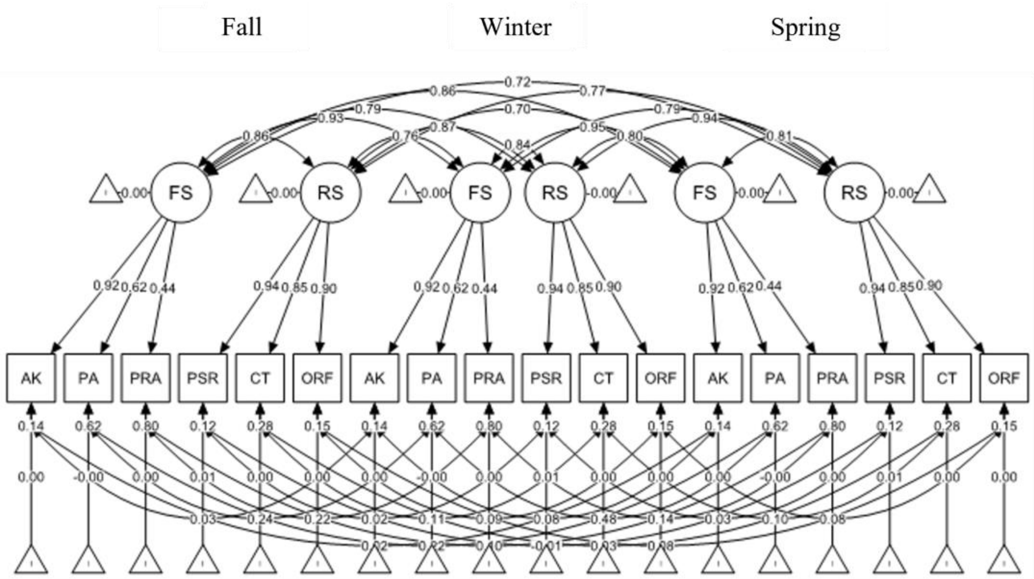
Table 6 displays the results for the incremental longitudinal invariance approach, the indices for each model, ANOVA for comparisons of fit between nested model invariance levels, and changes in fit indices.

Table 6
Robust fit indices of the longitudinal factorial invariance models of the IPAL.

Model	χ^2(gl)	RMSEA	CFI	TLI	ΔRMSEA	ΔCFI	Δ χ^2 (Δgl)	Pr(>χ^2)
Configural	435.12 (106)	0.057	0.982	0.973				
Metric	458.31 (114)	0.056	0.981	0.975	0.001	0.001	13.07 (8)	0.089
Scalar	458.48 (122)	0.053	0.981	0.977	0.002	0.000	0.17 (8)	1.000
Strict	532.55 (134)	0.048	0.983	0.981	0.004	0.002	8.55 (12)	0.740

The results show that the configural, metric, scalar, and strict models presented adequate fit indices (CFI < .98, TLI < .97, and RMSEA < .06). The reliability indices for each model were (1) configural model α = .953 and **Ω** = .930, (2) metric model α = .953 and **Ω** = .934, (3) scalar model α = .953 and **Ω** = .930, and (4) strict model α = .953 and **Ω** = .923. The results of the comparison of nested models through the Satorra–Bentler method (2001) for the difference in the χ2 statistic confirm the longitudinal invariance of the different forms of IPAL over the three measurement occasions. A comparison of the metric model with the configural model revealed that there were no statistically significant differences between the fits (Δχ2 (Δdf) = 13.07 (8), *p* = .08). Similarly, the contrast between the scalar and metric models confirms the level of scalar invariance, Δχ2 (Δdf) = .17 (8), *p* = 1.00. Finally, we confirmed the existence of strict invariance, Δχ2 (Δdf) = 8.55 (12), *p* = .74. A graphical representation of the longitudinal factorial invariance model can be found in Figure 4.

Figure 4
Longitudinal factorial invariance model for measuring IPAL Form A (fall), Form B (winter), and Form C (spring).



Note. The following abbreviations are used: FS: foundational skills, RS: reading skills, AK: alphabetic knowledge, PA: phonological awareness, PRA: print awareness, PSR: pseudoword reading, CT: cloze text, ORF: oral reading fluency.

Discussion

In this study, our approach centered on selecting skills identified by the NRP and NELP reports, supported by established reading models, to provide a comprehensive analysis of reading ability in first-grade primary school students. Contrary to our initial hypothesis of a single first-order factor, we uncovered a two-factor structure within the curriculum-based measurement (CBM) assessment, consisting of two distinct factors: foundational skills (FS) and reading skills (RS).

The identification of two latent factors, FS and RS, underscores the multidimensional nature of reading proficiency in early literacy development. FS encapsulates traditional foundational skills such as alphabetical knowledge, phonological awareness, and print awareness, closely aligned with the abilities highlighted by the NELP (2008). These precursor skills lay the groundwork for proficient reading, facilitating the transition from the logographic to the alphabetic and orthographic phases of literacy development (Frith, 1985). RS, on the other hand, corresponds to skills emphasized by the NRP (2000), including pseudoword reading, oral reading fluency, and comprehension.

A notable finding of this study is the relative stability of print awareness across time points compared with the more pronounced variation observed in other skills. This stability does not imply that print awareness lacks relevance in early reading assessment; rather, it highlights its foundational nature. Unlike skills such as alphabetic knowledge and phonological awareness, which are actively developed through instruction and show greater growth, print awareness is typically acquired earlier and remains more stable. This finding is consistent with previous findings suggesting that print awareness plays a crucial role in early literacy but may contribute less to literacy growth beyond the initial stages of reading acquisition (NELP, 2008).

The unexpected emergence of this two-factor structure suggests that while FS and RS are highly correlated, they maintain their unique identity within the broader construct of reading proficiency. Possible explanations for this finding could include the specific tasks included in the CBM assessment or other factors influencing reading development in first-grade students, warranting further investigation. Although FS and RS are highly correlated, model fit indices (e.g., RMSEA, CFI, TLI) from the confirmatory factor analysis indicate that a two-factor model provides a significantly better fit than a single-factor model does, supporting the distinction between foundational and reading skills. This suggests that while foundational and reading skills are interrelated, they remain conceptually distinct. Moreover, the two-factor structure aligns with theoretical models that distinguish early literacy precursors from conventional reading skills, supporting its validity as a framework for assessing reading development in first-grade students. This finding underscores the multidimensional nature of early reading proficiency and suggests that although these skills develop in parallel, they make distinct contributions to reading acquisition. These results align with established theoretical models of reading development, such as Frith's model (Frith, 1985) and the Simple View of Reading (Gough & Tunmer, 1986), which emphasize the importance of decoding fluency and comprehension in early literacy acquisition. Additionally, theoretical frameworks such as the dual-route cascaded model (Colheart et al., 2001) and the triangle model of reading (Plaut et al., 1996) suggest that decoding processes are fundamental in early reading development, highlighting the critical role of both foundational and higher-order reading skills in academic success.

The practical implications of our findings include informing educators and practitioners about the importance of early literacy instruction targeting fundamental components. By focusing on the development of alphabetical knowledge, phonological awareness, and print awareness, teachers can provide students with the necessary skills to become proficient readers. Additionally, the stability of the FS and RS factors over time, as evidenced by longitudinal factorial invariance, suggests that the CBM assessment IPAL is a suitable instrument for measuring reading ability in first grade and effectively monitoring student progress.

In summary, our study contributes to understanding the multidimensional nature of reading proficiency in first-grade students and underscores the importance of integrating both foundational and higher-order reading skills in early literacy instruction. Future research could explore additional factors influencing reading development and investigate the effectiveness of interventions targeting specific reading skills to improve overall reading outcomes. By building upon these findings, educators can enhance literacy instruction and support the academic success of all students.

Conclusions

This study demonstrated the longitudinal factorial invariance of a structure consisting of two latent variables: foundational skills (such as alphabetic knowledge, phonological awareness, and print awareness) and reading skills (including oral reading fluency, pseudoword reading, and comprehension) in the context of reading. This finding indicates that the IPAL, as a CBM, has a stable structure over time, facilitating meaningful comparisons in skill development across different points in the academic year. Importantly, all these skills are essential for reading (NRP, 2000; NERP, 2008). However, this study is not without its limitations. The selected sample consisted olely of students from the Canary Islands; thus, future research should encompass students from diverse regions of the country to ensure broader representation. Additionally, while the skills assessed through the IPAL approach are based on scientific research, it is imperative to acknowledge that they do not encompass all relevant skills. Hence, it would be beneficial to explore other abilities not addressed in this study. Another limitation pertains to the type of CBM measures employed, which are often considered specific-domain skills. While these skills are pivotal for reading proficiency, their development also hinges on general cognitive abilities such as working memory, executive functions, and attention (Compton et al., 2006). Consequently, future investigations should integrate measures of general-domain skills to yield a more comprehensive understanding of reading development. Future research should further examine the convergent validity of IPAL by comparing it to other curriculum-based assessments explicitly designed for Spanish-speaking students. Establishing convergent validity with well-established CBMs would strengthen its role as a reliable tool for progress monitoring and early identification of reading difficulties within Spanish-speaking populations. In Spain, there is a dearth of standardized CBM tools for the universal screening and monitoring of student learning progress in foundational learning areas. This study demonstrated the longitudinal factorial invariance of the IPAL Spanish version for first-grade students. Nonetheless, several areas of inquiry remain unexplored. For example, examining whether factorial invariance patterns persist across different educational contexts could validate the utility of CBM tests across various educational settings. Furthermore, investigating how interventions aimed at enhancing reading skills impact the invariant factorial structure over time could furnish valuable insights into intervention efficacy and its influence on underlying reading ability. Finally, a future research avenue should focus on similar studies across different languages to ascertain whether the factorial invariance identified in this study holds in diverse linguistic contexts, shedding light on similarities and discrepancies in reading development across languages.

Complementary information

Funding: This work has been made possible thanks to funding from the National R&D&i Plan of the Ministry of Economy and Competitiveness (Spain), with reference PID2019-108419RB-I00, with the PI as the first author.
Conflicts of interest/Competing interests: We have no known conflict of interest to disclose.
Acknowledgements: The authors would like to thank all the participant schools, teachers, students, and families.

References

Acha, J., Rodríguez, N., & Perea, M. (2023). The role of letter knowledge acquisition ability on children's decoding and word identification: Evidence from an artificial orthography. *Journal of Research in Reading, 46*, 358–375. https://doi.org/10.1111/1467-9817.12432

Álvarez-Cañizo, M., Suárez-Coalla, P. & Cuetos, F. (2018). Reading prosody development in Spanish children. *Read Writ, 31,* 35–52. https://doi.org/10.1007/s11145-017-9768-7

Baker, D.L., Alberto, P.C., Macaya, M.M., García, I., & Gutiérrez-Ortega, M. (2022). Relationship between the essential components of reading and reading comprehension in monolingual Spanish-speaking children: a meta-analysis. *Educ Psychol Rev, 34*, 2661–2696. https://doi.org/10.1007/s10648-022-09694-1

Baker, D. L., Cummings, K. D., Good, R. H. y Smolkowski, K. (2007). Indicadores dinámicos del éxito en la lectura (IDEL®): Summary of decision rules for intensive, strategic, and benchmark instructional recommendations in kindergarten through third grade (Technical Report No.1). Dynamic Measurement Group

Benjamin, R. G., & Schwanenflugel, P. J. (2010). Text complexity and oral reading prosody in young readers. *Reading Research Quarterly, 45*, 388–404. https://doi.org/10.1598/rrq.45.4.2

Braze, D., Katz, L., Magnuson, J.S., Mencl, W.E., Tabor, W., Van Dyke, J.A., Gong, T., Johns, C.L., & Shankweiler, D.P. (2016). Vocabulary does not complicate the simple view of reading. *Read Wri, 29,* 435-451. https:/doi.org/10.1007/s11145-015-9608-6

Calet, N., Gutierrez-Palma, N., & Defior, S. (2015). A cross-sectional study of fluency and reading comprehension in Spanish primary school children. *Journal of Research in Reading, 38*, 272–285. https://doi.org/10.1111/1467-9817.12019

Caravolas, M., Lervåg, A., Defior, S., Seidlová-Málková, G. & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science 24*(8), 1398-1407. https://doi.org/10.1177/0956797612473122

Catts, H. W., Nielsen, D. C., Bridges, M. S., Liu, Y. S., & Bontempo, D. E. (2015). Early identification of reading disabilities within an RTI framework. *Journal of Learning Disabilities, 48*(3), 281–297. https://doi.org/10.1177/0022219413498115

Child, D. (2006). *The essentials of factor analysis* (3rd ed.). Continuum.

Christ, T. J., Zopluoglu, C., Monaghen, B. D., & Van Norman, E. R. (2013). Curriculum-based measurement of oral reading: Multistudy evaluation of schedule, duration, and dataset quality on progress monitoring outcomes. *Journal of School Psychology, 51*(1), 19–57. https://doi.org/10.1016/j.jsp.2012.11.001

Clay, M. M. (1985). *The early detection of reading difficulties.* Heinemann.

Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98*, 394-409. https://doi.org/10.1037/0022-0663.98.2.394

Cuadro, A. & Trías, D. (2008). Desarrollo de la conciencia fonémica: Evaluación de un programa de intervención [Phonemic awareness development: An evaluation of an intervention program]. *Revista Argentina de Neuropsicología, 11,* 1-8. http://www.scielo.edu.uy/scielo.php?script=sci_arttext&pid=S1688-42212009000200006

Cuetos, F., Ruano, E., & Rodríguez, B. (2007). *PROLEC-R: Batería de evaluación de los procesos lectores* [PROLEC-R: Assessment battery for reading processes]. TEA Ediciones.

de la Calle, A. M., Guzmán-Simón, F., & García-Jiménez, E. (2018). Letter knowledge and learning sequence of graphemes in Spanish: Precursors of early reading. *Revista de Psicodidáctica, 23*(2), 128–136. https://doi.org/10.1016/j.psicod.2017.10.002

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, M. Coltheart & J. Marshall (Eds.), *Surface dyslexia* (pp. 301-330). Erlbaum.

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*(2), 117-135. https://doi.org/10.1016/j.jsp.2006.05.005

González-Valenzuela, M.J, Díaz-Giráldez, F., & López-Montiel, M.D. (2016). Cognitive predictors of word and pseudoword reading in Spanish first-grade children. *Front Psychol. 2016 May 31,* 7, 774. https://doi/org/10.3389/fpsyg.2016.00774

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *RASE: Remedial & Special Education, 7*(1), 6- 10. https://doi.org/10.1177/074193258600700104

Gough, P. B., Hoover, W. A. & Peterson, C. L. (1996). *Some observations on a simple view of reading.* Lawrence Erlbaum Associates Publishers.

Gutiérrez, N., Jiménez, J. E., & de León, S. C. (2022). Reading curriculum-based measures for universal screening in monolingual Spanish first graders. *Early Education and Development, 33*(6), 1036–1060. https://doi.org/10.1080/10409289.2021.1935537

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106*, 491–528. https://doi.org/10.1037/0033-295X.106.3.491

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662–720. https://doi.org/10.1037/0033-295X.111.3.662

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing. 2*, 127–160. https://doi.org/10.1007/BF00401799

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

International Dyslexia Association (2002). *Definition of Dyslexia.* International Dyslexia Association. https://dyslexiaida.org/definition-of-dyslexia/#

Jenkins, J. R., Schiller, E., Blackorby, J., Thayer, S. K., & Tilly, W. D. (2013). Responsiveness to intervention in reading: Architecture and practices. *Learning Disability Quarterly, 36*(1), 36–46. https://doi.org/10.1177/0731948712464963

Jiménez, J.E. (2019). *Modelo de respuesta a la intervención. Un enfoque preventivo para el abordaje de las dificultades específicas de aprendizaje* [Response to intervention (RTI) model: A preventive approach to addressing learning disabilities]. Ediciones Pirámide.

Jiménez, J.E., de León, S.C., & Gutiérrez, N. (2021). Piloting the response to intervention model in the Canary Islands: Prevention of reading and math learning disabilities. *The Spanish Journal of Psychology, 24,* e30. https://doi.org/10.1017/SJP.2021.25

Jiménez, J.E., & Gutiérrez, N. (2019). Indicadores de Progreso de Aprendizaje en Lectura [Indicators of basic early reading skills]. En J.E. Jiménez (Coord.), *Modelo de respuesta a la intervención. Un enfoque preventivo para el abordaje de las dificultades específicas de aprendizaje* [Material complementario] [Response to intervention (RTI) model: A preventive approach to addressing learning disabilities [Suplementary material]. Ediciones Pirámide.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). *semTools: Useful tools for structural equation modeling.* R package version 0.5-6. https://CRAN.R-project.org/package=semTools

Kim, Y. S., & Pallante, D. (2012). Predictors of reading skills for kindergartners and first grade students in Spanish: a longitudinal study. *Reading and Writing, 25,* 1-22. https://doi.org/10.1007/s11145-010-9244-0

Kline, R.B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford Press.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, *95*, 3–21. https://doi.org/10.1037/0022-0663.95.1.3

Lervåg, A., Bråten, I., & Hulme, C. (2009). The cognitive and linguistic foundations of early reading development: A Norwegian latent variable longitudinal study. *Developmental Psychology, 45*, 764–781. https://doi.org/10.1037/a0014132

Mariángel, S. V., & Jiménez, J. E. (2016). Desarrollo de la conciencia sintáctica y fonológica en niños chilenos: Un estudio transversal [Development of syntactic and phonological awareness in Chilean children: A

cross-sectional study]. *Revista Latinoamericana de Psicología, 48*(1), 1–7. https://doi.org/10.1016/j.rlp.2015.09.010
Milfont, T.L., & Fischer, R. (2010). Testing measurement invariance across groups: applications in cross-cultural research. *International Journal of Psychological Research, 3*, 111–130. https://doi.org/10.21500/20112084.857
Ministerio de Educación (2020). Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación [Organic Law 3/2020, of December 29, amending Organic Law 2/2006, of May 3, on Education]. *Boletín Oficial de Estado, 340*, 122868-122953. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-17264
Nation, P. (2015). Principles guiding vocabulary learning through extensive reading. *Read. Foreign Lang. 27*, 136–145.
National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel.* Washington, DC: National Institute for Literacy.
National Reading Panel (NRP) (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction:* Reports of the subgroups (NIH Publication No. 00-4754). U.S. Government Printing Office.
Newsom, J.T. (2015). *Longitudinal structural equation modeling: A Comprehensive introduction* (1st ed.). Routledge. https://doi.org/10.4324/9781315871318
O'Brien, E.M. (2008). *From beginning to stellar: Five tips on developing skillful readers*. Center for Public Education.
Papadopoulos, I., & Bourogianni, M.E. (2024). Delving into word and print awareness in 4-year-old children. *British Journal of Education 12*(3):41-54. https://doi.org/10.37745/bje.2013/vol12n34154
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasiregular domains. *Psychological Review, 103*, 56–115. https://doi.org/10.1037/0033-295X.103.1.56
Puranik, C.S., Lonigan, C.J, & Kim, Y.S. (2011). Contributions of emergent literacy skills to name writing, letter writing, and spelling in preschool children. *Early Child Res Q, 26*, 465-474. https://doi.org/10.1016/j.ecresq.2011.03.002.
Rehfeld, D.M., Kirkpatrick, M., O'Guinn, N., & Renbarger, R. (2022). A meta-analysis of phonemic awareness instruction provided to children suspected of having a Reading Disability. *Lang Speech Hear Serv Sch*; *53*, 1177-1201. https://doi.org/10.1044/2022_LSHSS-21-00160.
Revelle, W. (2023). *psych: Procedures for psychological, psychometric, and personality research.* Northwestern University, Evanston, Illinois. R package version 2.3.6. https://CRAN.R-project.org/package=psych
Rice, M., Erbeli, F., Thompson, C.G., Sallese, M.R., & Fogarty, M. (2022). Phonemic Awareness: A meta-analysis for planning effective instruction. *Reading Research Quarterly, 57*, 1259-1289. https://doi.org/10.1002/rrq.473.
R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for statistical computing, Vienna, Austria. https://www.R-project.org/
Rodríguez, C., Gutiérrez, N., Seoane, R. C., González, D., & de León, S. C. (2024). Reading Spanish prosody: The role of word reading and syntactic knowledge. *Journal of Literacy Research*. https://doi.org/10.1177/1086296X241244705
Rosseel, Y. (2012). "lavaan: An R Package for structural equation modeling." *Journal of Statistical Software, 48*, 1–36. https://doi.org/10.18637/jss.v048.i02
Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514. https://doi.org/10.1007/BF02296192
Sellés, P. N., Martínez, T. G., & Vidal-Abarca, E. (2010). Batería de Inicio a la Lectura (BIL 3-6): Diseño y características psicométricas [Initial reading battery (BIL 3–6): Design and psychometric characteristics]. *Bordón. Revista de Pedagogía, 62*(2), 137–160.
Shinn, M. R., Rosenfield, S., & Knutson, N. (1989). Curriculum-based assessment: A comparison of models. *School Psychology Review, 18*, 299-316. https://doi.org/10.1080/02796015.1989.12085428
Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. & Stahl, S. A. (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational Psychology, 96*, 119–129. https://doi.org/10.1037/0022-0663.96.1.119.
Serrano, F., Defior, S. & Jiménez, G. (2005). *Evolución de la relación entre conciencia fonológica y lenguaje escrito en niños españoles de primer curso de Educación Primaria* [Development of the relationship between phonological awareness and written language in first-grade Spanish children]. Libro de actas del II congreso Hispano Portugués de Psicología: Coruña.
Snow, C. (2002). *Reading for understanding: Toward an R & D program in reading comprehension*. Santa Monica, CA: RAND Corporation.
Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407. https://doi.org/10.1598/RRQ.21.4.1
Stecker, P. M., Lembke, E. S., & Foegen, A. (2008). Using progress-monitoring data to improve instructional decision making. *Preventing school failure: Alternative education for children and youth, 52*, 48-58. https://doi.org/10.3200/PSFL.52.2.48-58
Torppa, M., Georgiou, G. K., Lerkkanen, M. K., Niemi, P., Poikkeus, A. M., & Nurmi, J. E. (2016). Examining the simple view of reading in a transparent orthography: A longitudinal study from kindergarten to grade 3. *Merrill-Palmer Quarterly*, *62*, 179–206. https://doi.org/10.13110/merrpalmquar1982.62.2.0179
Vázquez, M., Laakman, A. L., Loria Garro, E. S., Tan, S. X. L., & Keller-Margulis, M. A. (2024). Curriculum-based measurement in languages other than English: A scoping review and call for research. *Contemporary School Psychology, 28*, 341–352. https://doi.org/10.1007/s40688-023-00487-z
Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2016). Bidirectional relations between text reading prosody and reading comprehension in the upper primary school grades: A longitudinal perspective. *Scientific Studies of Reading, 20*, 189–202. https://doi.org/10.1080/10888438.2015.1128939