

Research Article

Can We Identify Struggling Writers in the Classroom Using Pseudowords?

Mariona Pascual,^a  Mario Figueroa,^b  and Naymé Salas^c 

^aDepartament de Pedagogia Aplicada i Psicologia de l'Educació, Institut de Recerca i Innovació Educativa, Universitat de les Illes Balears, Mallorca, Spain ^bDepartament de Psicologia Bàsica, Evolutiva i de l'Educació, Universitat Autònoma de Barcelona, Spain ^cDepartament de Didàctica de la Llengua i la Literatura, i de les Ciències Socials, Universitat Autònoma de Barcelona, Spain

ARTICLE INFO**Article History:**

Received September 7, 2025

Revision received January 13, 2026

Accepted June 22, 2026

Editor-in-Chief: Karla N. Washington

Editor: Benjamin Munson

https://doi.org/10.1044/2026_LSHSS-25-00170**ABSTRACT**

Purpose: Early identification of literacy difficulties is essential to implement effective remediation strategies. Often, tests are time-consuming and may require expert practitioners to administer or interpret them. This study tested an easy-to-use spelling tool screen for early writing difficulties. Specifically, we explored whether a pseudoword spelling task was effective for the early identification of struggling writers (SWs). We examined its discriminatory ability and differences in performance as a function of item features.

Method: Forty-eight primary SWs and 60 typically developing (TD) peers completed a pseudoword dictation task in Catalan (Pseudoword Spelling Evaluation in Catalan Language [PSE-CL]). We analyzed key psychometric properties of the task, including its reliability and discriminatory ability. Item manipulation allowed for an analysis between and within group performance based on item features (i.e., length and frequency of the original base word).

Results: The task showed good internal consistency and reliability, with grade-specific variations. It effectively distinguished SWs from TD children (area under the curve: 0.87), achieving 79% accuracy at the cutoff point. A cross-classified multilevel model examining effects within items and between participants showed that SWs had substantially lower spelling accuracy than TD peers. Specifically, their predicted likelihood of spelling a word correctly was about seven times lower. The model also showed a significant Group $\times$ Grade interaction, indicating that the difference between SWs and TD peers was especially pronounced in Grade 3.

Conclusions: The PSE-CL is an easy-to-administer tool for speech therapists and teachers to identify SWs within the classroom. Children who struggle with writing appear to have impaired phonological-mapping abilities, especially in third grade. The theoretical and practical educational implications are discussed.

Regardless of whether a child has a neurological or developmental disorder, some children consistently exhibit difficulties in one or more dimensions of writing, such as text quality, organization, or spelling. Recent meta-analytic studies have examined the writing performance of children who may struggle (hereafter, struggling writers [SWs]) in these areas, identifying spelling as a particularly vulnerable domain where students often perform below their peers

(Troia, 2025). These challenges are observed across diverse populations, including those with learning disabilities, reading difficulties, developmental language disorder, attention-deficit/hyperactivity disorder, and autism spectrum disorder. Typically, identification of writing difficulties relies on standardized assessments (Dockrell et al., 2019) and curriculum-based measurement (CBM) tools (Piercy et al., 2023), which emphasize the importance of evaluating writing as a multi-dimensional construct and adopting performance-based approaches rather than depending solely on external diagnoses that may overlook written expression (Cragg & Nation, 2006; Williams & Larkin, 2013). This perspective

Correspondence to Mariona Pascual: mariona.pascual@uib.cat.

Disclosure: The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

frames writing difficulties as production based, highlighting not only low attainment on specific tasks but also broader impacts on sentence generation and overall text quality (Balade et al., 2024; Dockrell et al., 2018). Importantly, SWs often display early risk indicators: Their difficulties are not merely secondary manifestations of a main neurocognitive impairment that are common across different diagnostics but also include learners who remain undetected by traditional screening frameworks or have no access to appropriate screening tools, even though they require tailored interventions (Dunn, 2021; Harris et al., 2006; Helsel & Greenberg, 2007).

Interpreting performance in SWs can shed light on the missing links, where children who do not fit a diagnosis might still require assessment and intervention to reach their peers' developmental advances (Arfè & Dockrell, 2023; Pascual & Salas, 2021). It should be noted that SWs often receive their intervention in public or nonprofit organizations, so teachers or speech therapists in schools may lack the resources and time to design analytical tools that provide a detailed analysis of difficulties (Daffern & Ramful, 2020; Denman et al., 2021; Dockrell & Marshall, 2015). There is, therefore, a need for the development of a tool that enables the detection of difficulties in spelling in a prompt and efficient manner for teachers, which in turn meets the sustainable development goals (SDGs): a set of 17 global objectives established by the United Nations to promote inclusive and equitable quality education, among other aims. Early identification of spelling difficulties aligns particularly with SDG 4 (Quality Education) while also supporting broader aims such as reducing educational disparities (SDG 10) and contributing to students' overall well-being (SDG 3) through timely and appropriate interventions (Nedungadi et al., 2024). In this article, we present a spelling screening tool for children in the middle grades of primary education.

Spelling Skills

Writing difficulties are typically considered a by-product of other, primary learning and language disorders. As a result, they are often underdiagnosed and undertreated (Coker et al., 2018; Dockrell et al., 2019, 2022; Pascual & Salas, 2021). Spelling screening tools can be really useful, because they tap complex phonological skills, which are at the core of literacy difficulties (Wiley et al., 2024). Thus, spelling tasks can shed more light on atypical spelling development trajectories, as well as the processes underlying difficulties with other phonological tasks. Besides contributing to advancing our knowledge of literacy difficulties and development, spelling tasks may also serve a more practical need. Teachers and practitioners have expressed the need for easy-to-use resources to identify children at risk of literacy difficulties, as well as those who could benefit from intensive support or from being referred to more targeted

clinical intervention (Graham et al., 2018). It is critical to have resources for both clinical and school settings to detect literacy difficulties beyond the early primary grades. This is because children must be able to face literacy demands in higher grades, where automatization of decoding and spelling abilities is assumed to be established (Fitzgerald & Shanahan, 2000).

Spelling development depends on individual skills, such as letter knowledge and phonological awareness (i.e., identifying and manipulating phonemes; Caravolas et al., 2012; Kim et al., 2010; Nielsen & Juul, 2016; Verhagen et al., 2010). These core skills set the foundation for later development of higher order writing skills (Berninger & Winn, 2006), which, in turn, become powerful predictors of school achievement and essential for literacy development (Caravolas et al., 2001; Graham & Santangelo, 2014; Treiman et al., 2019).

Basic phonological skills, which are usually targeted and acquired during the early school years, are key to efficiently and precisely encoding and developing lexical representations (Landi & Perfetti, 2007; Perfetti & Hart, 2002; see lexical quality theory; Perfetti, 2007). This development is not restricted to schooling ages, as the lexical representations can still play a pivotal role in the orthographic precision abilities of adults (e.g., in reading tasks; Adelman et al., 2014; Meade et al., 2018), as spelling ability depends on many factors, ranging from vocabulary to linguistic abilities or even literacy exposure (Perfetti, 2007).

Thus, it is not surprising that, in early childhood, basic phonological skills are pivotal (Goswami, 2001; Melby-Lervåg et al., 2012). Core phonological knowledge (McLeod & McCormack, 2015), such as phonological and phonemic awareness, is learned in the first years of schooling (Caravolas et al., 2001, 2005), but spelling does not solely build upon grapheme-phoneme mapping (Seidenberg et al., 2022; Treiman, 2018), and thus, the development of spelling skills requires further refinement during primary education. For example, children also rely on graphotactical information (e.g., contextual features, letter patterns; Treiman et al., 2019; Vale & Perpétua, 2020) or morphological, lexical, and orthographic knowledge that develop in parallel (Caravolas et al., 2005; Daffern et al., 2015; Salas, 2020; Treiman et al., 2019; see the triple word form theory; Bahr, 2015). Therefore, phonological abilities seem to be progressively relegated from the spotlight, making way for other strategies that allow students to regularize and perfect spelling abilities (Alves et al., 2018; Bear et al., 2012; Treiman, 2017; Treiman et al., 2002).

However, spelling trajectories in SWs may follow a different pattern, with slower and more irregular development (e.g., Arfe et al., 2014; Dockrell et al., 2014; Graham et al., 2018). In these populations, difficulties may be linked to basic phonological deficits as they have been found to be

related to performance on spelling tasks across different groups with different diagnoses (Hoefflin & Franck, 2005; Larkin & Snowling, 2008; Silliman et al., 2018). For example, deaf children tend to produce plausible phonological errors, and those with low aided audibility produce significantly more phonological omissions (Quick et al., 2021), whereas phonological errors in children with developmental language disorder consist largely of consonant omissions (e.g., “ecept” for “except”) and vowel and consonant substitutions (e.g., “dack” for “duck” or “den” for “then,” respectively; Joye et al., 2020). To this end, the current study focuses on developing a pseudoword screening tool to discriminate between children with writing difficulties and those with typical development.

What Does a Pseudoword Spelling Task Involve?

Pseudowords do not have a semantic or lexical representation (de Bree & van den Boer, 2019), so they offer a window into the exploration of phonological spelling abilities, in a context that can be exempted from morphological, lexical, and semantical knowledge (i.e., there is no abstract representation or a prior full lexical correspondence knowledge to retrieve). In other words, the errors observed in the task are proxies for poor phoneme-to-grapheme mapping (Andrews et al., 2020; Daffern & Ramful, 2020). Errors in dictation tasks have been used for screening purposes (Bear et al., 2012; Daffern & Ramful, 2020; Sacre & Masterson, 2000), and especially, performance in pseudoword tasks has been used as a proxy to detect, for example, writing difficulties in children with dyslexia (Juul & Petersen, 2017; Rothe et al., 2015; Van Vreckem & Desoete, 2018).

When exploring the performance of children who might struggle with writing in pseudoword spelling tasks, conflicting results have arisen. While several studies found impaired performances in pseudoword accuracy (e.g., Friend & Olson, 2008; Juul & Petersen, 2017; Rothe et al., 2015; Van Vreckem & Desoete, 2018), others did not find any difference between children at risk of writing difficulties and their peers (Bourassa & Treiman, 2008; Cassar et al., 2005; Diamanti et al., 2014; Protopapas et al., 2013). For instance, Dutch-speaking children with dyslexia exhibited sustained phonological spelling difficulties through to Grade 6 (Van Vreckem & Desoete, 2018) and Grade 9 (Juul & Petersen, 2017). Interestingly, besides intrinsic factors that can hinder performance, there are two within-item factors that can influence performance in pseudowords: stimulus length (Juul & Petersen, 2017) and the frequency of the word from which the pseudowords have been derived (Grainger et al., 2000; Ozubko & Joordens, 2011; Suárez-Coalla et al., 2016). As per the dual-route model (Angelelli et al., 2018; Cataldo & Ellis, 1988; Martin et al., 2020;

C. Perry et al., 2002), pseudowords have to be approached from a sublexical pathway, and thus, manipulation of the item features would be detrimental to better understand the sublexical constraints.

Our Study

Our aim was to investigate whether a pseudoword spelling task targeting core phonological encoding skills helps distinguish between at-risk and typical students in Catalan, an intermediate-depth language (Share, 2008), based on previous findings of the prevalence of phonologically based struggles at these ages (Juul & Petersen, 2017). In this case, Catalan, as an intermediate-depth language, has eight vocalic and 33 consonantal sounds, which can be represented by different graphemes depending on orthographical, morphological, lexical, or phonological criteria (Salas, 2020). Catalan is a Romance language closely related to Spanish, with which it shares a high degree of morphosyntactic similarity (Salas, 2020; Tolchinsky, 2021). Despite this, there is a paucity of evidence on spelling difficulties in Catalan, as well as a lack of tools to detect them. The few studies conducted on Catalan children show that phonographic spelling is mastered earlier than orthographic and lexical spelling (Llaurado & Tolchinsky, 2016; Salas, 2020; Tolchinsky, 2021). Only 10% of errors or less are phonographically based at the third school grade (Tolchinsky, 2021). However, SWs may exhibit poorer performance and a higher error rate on these words than typically developing (TD) children, who are already proficient in the early elementary grades.

To test the psychometric properties of this task and determine whether phonological encoding skills still help distinguish between at-risk and TD children, we first conducted a task-reliability analysis with TD children and SWs, assessing its construct validity and discriminative properties. Second, to assess the task in depth, we examined each group's performance based on the item's properties. Beyond assessing task potential, our main aim is to explore performance within groups and pinpoint the main effects that might be conditioning performance in the task (i.e., stimulus length or the frequency of the base word).

Accordingly, this study addresses the following research questions: Can a pseudoword spelling task, which targets phonological encoding skills, reliably distinguish between at-risk SWs and TD Catalan-speaking children? Do item-level properties, such as stimulus length and base-word frequency, influence task performance?

We hypothesized that SWs would demonstrate significantly lower accuracy than TD peers, indicating persistent weaknesses in phonological encoding skills. Longer pseudowords were expected to reduce accuracy for all participants,

with a more pronounced effect in the SW group. Additionally, a facilitatory effect of base-word frequency was anticipated, resulting in better performance on pseudo-words derived from higher frequency words. If frequency effects are observed, this would suggest the influence of pre-established, contextually structured spelling patterns over phonological units (Deacon et al., 2008; Treiman et al., 2002).

Method

Ethics Statement

This project was carried out in compliance with the guidelines of the research ethics committee of the Universitat Autònoma de Barcelona. As a part of a larger project, the ethics committee approved the large-scale project (January 22, 2016). As per the committee's usual procedure at that time, no reference number was provided with this signed approval. Complementarily, for this specific data collection, a good-practices document was signed attesting that all data were collected in accordance with the principles and guidelines established by the committee. This project was reviewed annually by an external committee that specifically evaluated compliance with ethical guidelines, adherence to the project's aims, and advancement of core objectives. Written informed consent to participate in this study was provided by the participants' parents or legal guardians.

Participants

We collected data from five urban schools in Barcelona (Catalonia, Spain), which were participating in a larger intervention study. All participating schools were public schools located in middle-to-high socioeconomic status (SES) areas, for a relatively homogeneous average of the city's educational and occupational representation. We selected "regular" schools (excluding those labeled by the Catalan Department of Education as "vulnerable" or "of high complexity"), as previous research has shown how the SES profile of a school can affect literacy development beyond the familial SES (Salas & Pascual, 2023). Beyond the classification of "regular" schools by governmental agencies, we obtained information on parents' occupations and quantified it using Ganzeboom et al.'s (1992) index, which ranges from 20 to 90. Participants scored, on average, 57.94 points (range: 30.50–74.80), which is linked to a generalized middle SES (see Table 1 for a breakdown by group). In all settings, the main language of instruction was Catalan, and the majority of children had learned Catalan in the public school by no later than 6 years of age (first grade of primary education). Almost all students

reported being enrolled since kindergarten in the public system.¹

A total of 98 students from a sample² of 107 were included in this study, and full demographic information is shown on Table 1. Participants were enrolled in Grades 3 and 5. To identify children at risk of writing difficulties, students completed a series of writing tasks 1 year before the data collection for the present study. To be classified as an SW, children had to perform at least 1 *SD* below the mean for their grade across (a) two handwriting fluency tasks; (b) two text composition tasks, namely, a narrative and an argumentative essay; and (c) a phonological-spelling task. Children at risk of writing difficulties were thus those who either showed low handwriting fluency or low productivity in text composition (fewer words produced) or failed to reproduce the phonological structure of dictated words, of varying lengths and complexity. In addition, to be categorized as an SW for the present study, writing difficulties in either domain (handwriting, productivity, or spelling) had to be confirmed 16 months afterward (Pascual et al., 2020; Pascual & Salas, 2021; for a detailed report of performance on the screening tasks, see Table A1 of the Appendix). This identification protocol was developed following studies targeting struggling profiles in the absence of standardized tasks (Mahfouz & Mohamed, 2023; Snyder & Ayres, 2020). Our screening procedure targeted students whether or not they had received any kind of learning or language diagnosis and regardless of whether they had been receiving some curricular adaptation (Pascual & Salas, 2021). A total of 43 children fulfilled the criteria and constituted our SW group. Independently of the CBM measures, 11 SW participants had also been priorly formally diagnosed with a neurodevelopmental disorder at some point during schooling and had an individualized educational program implemented by the school when they identified at risk of language delay behaviors. An age-matched, TD group of students was randomly drawn from the larger sample as long as they had (a) no prior language or learning disorder and showed within-average performance across the abovementioned writing tasks (i.e., within 1 *SD* in handwriting, spelling, and text-generation abilities). A total of 55 children constituted the TD group. Participants' distribution per group was balanced across grades, $\chi^2(1) = 0.508$, $p = .476$, as well

¹Kindergarten (ranging from 3 to 6 years old) is not mandatory in Spain and might be carried out in other settings (e.g., private, rural schools, caretakers within the family). Two participants, one of each group, reported accessing the public school system at first grade.

²Nine participants were excluded because they had received a diagnosis of high abilities or "gifted" ($n = 1$), severe intellectual disability ($n = 1$), and behavioral or emotional disorder ($n = 5$). One student was excluded due to its newcomer status ($n = 1$); and another one, because they did not complete the task ($n = 1$).

Table 1. Demographic information per group.

Group	Grade	Sex	Age			ISEI		
	3rd; 5th (n)	Boys; girls (n)	M	SD	Range	M	SD	Range
Struggling writers (n = 43)	25; 18	33; 10	114.5	13.18	85–135	53.56	13.76	30–75
Typically developing controls (n = 55)	28; 27	29; 26	105.80	16.58	81–136	61.64	10.93	32–75

Note. ISEI = International Socioeconomic Index.

as the participant's sex distribution per grade, $\chi^2(1) = 0.039$, $p = .844$.

Task Development

The task of Pseudoword Spelling Evaluation in Catalan Language (PSE-CL) was constructed with pseudowords adapted from the Procesos Lectores (PROLEC; Cueto et al., 2010) task of word reading, by substituting between one and four graphemes in each word. PROLEC stands for “PROcesos LECTores” and is a standardized battery designed to assess reading processes in Catalan-speaking children (Cueto et al., 2010). It evaluates components such as word recognition, reading comprehension, and phonological processing, making it a widely used tool in educational and clinical contexts. All items were phonotactical modifications of the real Catalan words included in PROLEC. These words were chosen after a discussion among a team of experienced researchers with expertise in linguistics, language didactics, and clinical training in speech therapy. Items were selected considering two main factors: frequency³ of the base word (low–high frequency) and length of the stimuli (short words of one to two syllables or long words of three to four syllables). Accordingly, and due to the target of the task, pseudowords were created to account for both phonological and phonetic rules of occidental Catalan (see Table 2 for a summary and examples of items), such as representative syllabic structures of Catalan (V, VC, VCC, CV, CVC, CVCC, CCVCC), as well as frequent phonetic events either within syllables or between syllables, such as those produced due to chronolectal and geolectal variation (e.g., vocalic reduction of [o] to [u], which does not happen to all dialects) to facilitate articulatory structures or simplification of consonant groups (Bonet & Lloret, 1998; Carrera-Sabaté et al., 2014; Solà et al., 2008; Wheeler, 2005).

We first carried out a pilot study to test the pseudoword's reliability on an unselected sample of the same

demographic features with 48 third and fifth graders ($n = 24$ in each group). The first version of this test consisted of 40 items. Once piloted, we calculated the internal consistency through Cronbach's alpha. Cronbach's α coefficient was .57. Due to low reliability, an exploration of the task's items prompted modifying six of the items. Modifications were due to (a) ceiling effects on accuracy ($n = 3$), (b) high similarity with a real word that led to sustained lexicalizations ($n = 1$), and (c) increased variability of syllabic structures ($n = 2$). Four additional items were incorporated in the final task to increase the utterances of articulatory events to ensure variability.

Procedure

Data gathering was carried out in quiet classrooms during school hours and within the school premises with small groups of a maximum of five children. This task was implemented as part of a larger longitudinal project; thus, it was administered alongside other written tasks at the end of the school year (May to June) in an approximately 40-min-long session (task implementation ranged from 12 to 14 min). Before starting, the experimenter explained to participants that they had to write down the words they would hear. The instructions specified that the items would be unknown words to them, but they had to try to write them as well as possible. The experimenter, a Catalan native speaker, enunciated the words twice, and students had to write them by hand. Enunciation was maintained at a steady pace, with sufficient time in between to write, but it was indicated to participants that, if they did not have time to complete the word before hearing the next, they had to jump to the next one.

Data Analysis

All data processing was carried out with R (R Core Team, 2020). Participants' data were transcribed by two research assistants, who typed all responses, preserving spelling from the original productions, and a computational analysis was developed to avoid bias from raters. Next, we needed to generate a glossary of all phonologically plausible versions for each item, accounting only for grapheme–phoneme

³Frequency was determined by the online search engine of the Catalan frequency dictionary of the Institut d'Estudis Catalans (IEC): <https://latel.upf.edu/terminotica/diccfreq.htm> (Rafel, 1996), which draws data from the corpus of the IEC (29.266.353 words extracted from 2.288 non-literary texts).

Table 2. Summary of main Catalan vocalic and consonant events that were considered within word.

Phonological event	Variable	Real-word example	PSE-CL item example
Vowel reduction in unstressed syllables	a/e > [ə] o/u > [u]	oliva [uliβə] <i>olive</i>	[bəkurdɔl]
Diphthong	ascending [jə] descending [áj]/[éj]	dia [djə] <i>day</i> lleï [léj] <i>law</i>	[trənsjə]
Epenthesis	[ə]-	escàner [əskàner] <i>scanner</i>	[əskríə]
Sonorization	tg → [dʒ]	fetge [fédʒə] <i>liver</i>	[préɲdʒə]
Assimilation	[mf], [nt], [ng]	amfibi [əmʃibi] <i>amphibian</i>	[trumfèʒa]
Gemination	[bb]	futbol [fubbol] <i>football</i>	[siobbə]
Deafening at the end of the word	[b] > [p] [d] > [t] [g] > [k]	verd [bért] <i>green</i>	[bork]
Mutening at the end of words	-r -t	fuster [fusté] <i>carpenter</i>	[munəló]
Simplification of consonantic groups	“-ld,” “-lt,” “-nd,” “-nt” “-m(p)t-” “-m(p)c”	profund [prufún] <i>deep</i> compte [komtə] <i>account</i>	[ufpərin]

Note. PSE-CL = Pseudoword Spelling Evaluation in Catalan Language.

correspondences. We developed a script with the packages *stringi* (Gagolewski, 2022) and *matchmaker* (Kamvar, 2020) for regular expressions (Wickham et al., 2023). All letter alternations corresponding to contrasting phonemes were deemed incorrect. For example, “s” is accepted for both the voiced [z] and voiceless [s] phonemes, but the grapheme “z” would only be accepted for the sound [z], and the graphemes “ss” and “ç” would only be correct for [s]. Overregulatory patterns derived from orthographic knowledge were accepted, such as adding “h” to words starting with vowels or adding an accent to stressed syllables. On the other hand, writing graphemes from the Spanish alphabet, such as “ch” and “ñ,” was not considered a correct representation. The item’s structure was regularized to account for present or absent clusters (e.g., “h”), anchored letters that did not allow for variation, and clustered allographic variations. We analyzed our glossary items. All words were transformed to regular expression syntax, which means that different alternations within the word were accepted and deemed as correct. These patterns were tested and debugged, scanning for errors, false positives, and false negatives and for over- and underinclusion with the web-based app *Regex* (regex101.com; Firas Dib, 2021). The automated scoring was implemented using a dictionary-based function that matches results to entries in the glossary via a key search, awarding 1 point for each correctly spelled word. The first author and two research assistants visually inspected all the results, scanning for overinclusion or false positives. Before scoring words with

errors as incorrect (0), they were further inspected by the research assistants to ensure no false negatives.

Psychometric properties were examined using the package *TAM* (Robitzsch et al., 2024). We report Cronbach’s alpha results, but to expand on these results, given critical reviews of the usage of Cronbach’s alpha (Hayes & Coutts, 2020; Revelle & Condon, 2019), we accounted for the task performance with McDonald’s omega (ω ; McDonald, 1999), which does not require all items to contribute equally to the measured construct. We also performed a second set of analyses to assess the discrimination and difficulty of the different items using item response theory modeling (Choi & Asilkalkan, 2019), such as expected a posteriori (EAP) reliability, which is a model-based index that reflects observed score variance rather than measurement error. These reliability measures have been shown to be reliable at small scales and have previously been used in spelling tasks (Daffern & Ramful, 2020).

To assess how well the task distinguishes the SW group from the TD group, we conducted a receiver operating characteristic (ROC) analysis and further analyzed the results. The ROC curve plots true-positive rates versus false-positive rates across different thresholds, and the area under the curve (AUC) measures discriminability, which is how effectively this task can differentiate SWs across the range of scores. ROC curves have been used to assess classification properties (Gil et al., 2021; Giovagnoli et al., 2020; Ritchey & Coker, 2014). To find

the optimal cutoff points for the task, using a reliable measure independent of group size, we further explored the Youden index (J) with the package `cutpointr` (Thiele & Hirschfeld, 2021). Importantly, the ROC curve analysis was also performed by grade to determine whether its discriminatory effect differs, given the crucial turn point between the third and fifth grade.

Afterward, we ran cross-classified multilevel models (CCMLMs) in Mplus (Version 8; Muthén & Muthén, 2025). CCMLM is used for nested structures, such as in this task, to account for the inherent cross-classification of the data, given that there are multiple participant answers per word and multiple words answered by participants, and thus, random effects are accounted for participants and items simultaneously (Van den Noortgate et al., 2003). To explore the performance in the task by group and grade while integrating both higher levels of variation (spelling accuracy by students and by words), we built a cross-classified analytical framework, also known as multiple-membership analysis, which assessed the two higher level sources of variation simultaneously, namely, between children and between word variance (Cafri et al., 2015). For this analysis, we used the Mplus (Version 8.3) cross-classified procedure (Muthén & Muthén, 2025). The intraclass correlation coefficients (ICCs) were calculated for children and words using the 2/3 formula for the measurement-level variance for binary responses (Wu et al., 2012).

A preliminary test for possible school differences (fixed effect) in terms of accuracy performance at the student's level was conducted within this analytical framework and found all comparisons between them to be non-significant (all b s < 1; all p s $\geq$.1). Therefore, school was not considered as a factor in all subsequent analyses. This insignificant level remained across all other pairwise comparisons. Thus, school affiliation was excluded from further analyses, given that schools were not a factor in explaining word accuracy difference. A similar test was applied to test grade difference and found that children at both grades performed at a similar accuracy of spelling, yet in the final model, we included grade and its interaction term with the group affiliation, that is, diagnosed versus typically developed children. It is important to note that five schools were not enough to include as a higher random level. Complementarily, between student ($n = 108$) and word ($n = 44$) variations showed ICCs higher than the threshold for integrating these sources of variation ($ICC_{\text{student}} = .068$, $ICC_{\text{word}} = .067$). In the final model, we included word length and frequency as additional independent indicators at the word level as well as group and grade at the student level. An interaction term between group and grade was tested in the final model. Predicted marginal means were calculated using the Mplus Model Constraint option.

Results

Psychometric Properties of the Pseudoword Task

Internal consistency for the final 44 pseudowords was assessed, with an overall acceptable Cronbach's α of .73. Assumption of unidimensionality was met, as omega statistics for the proportional variance of all common factors was above .7 ($\omega = .75$; Nahm, 2022; Ten Berge & Sočan, 2004; Revelle & Condon, 2019). With an EAP reliability of .673, the mean item difficulty was $\sigma = -1.80$ ($SD = .90$). In summary, the task showed acceptable internal consistency and reliability.

The AUC was 0.86 ($SE = 0.03$; $p < .001$, 95% confidence interval [CI] [0.80, 0.94]), indicating a fair performance of the task to differentiate between SWs and TD children, with a Youden index ($J = .575$; Youden, 1950) that signaled 79% accuracy as the cutoff point, indicating that scores below this threshold would fairly distinguish SWs from their peers. Nonetheless, when observed by grade, we found an excellent diagnostic performance, with an AUC of 0.96 in third grade ($J = .78$; optimal cutoff point at 82%), with sensitivity and specificity of .96 and .82, respectively, but for fifth grade, there is a decrease to a discriminative value with an AUC of 0.69 with sensitivity and specificity of 0.81 and 0.61, respectively ($J = .43$; optimal cutoff point at 79%).

Effects of Group and Grade in the PSE-CL

In the TD group, the M_{accuracy} was 0.89 ($SD = 0.05$) for Grade 3 and 0.85 ($SD = 0.05$) for Grade 5. In the SW group, the M_{accuracy} was 0.71 ($SD = 0.10$) in Grade 3 and 0.79 ($SD = 0.09$) in Grade 5. The cross-classified modeling main effect results indicated that the SW group's performance was significantly below that of the control group ($b = -0.505$, posterior standard deviation (PSD) = 0.069, 95% CI [-0.642, -0.371], $p < .001$). In other words, when we transformed the exponent of the b coefficient to obtain the odds ratio, the mean predicted probability of success for an SW participant was seven times lower than that of the control group participant. However, the length and frequency effects at the word level were found to be insignificant; that is, no difference in the probability of accurate spelling was observed between longer and shorter words or between more and less frequent words. Overall, the model fit quality was high (posterior predictive probability = .489; 95% CI for c2 difference: [-11.98, 13.59]). Next, we added the grade level as an additional main effect at the student level, yet we could not reject the null hypothesis that grade was not a factor in the probability of performing accurate spelling. An interaction term between group and grade was added and was found to be significant ($b = 0.461$, PSD = 0.132,

95% CI [0.205, 0.722], $p < .001$). By ranking the four subgroups (SW fifth graders, TD fifth graders, SW third graders, and TD third graders), we found that typical students at fifth and third grades showed the highest probability to perform an accurate spelling (probability [prob.] = .768 and .801, respectively), whereas the probability of accurate spelling for SW third graders was the lowest (prob. = .664), and that of the SW fifth graders was slightly higher (prob. = .721). These predicted probabilities are presented in Figure 1. By comparing the predicted probabilities, we show that, contrary to our intuition, the TD third graders had the highest average probability of successful spelling, while the SW third graders were the lowest, and the SW fifth graders were the second lowest.

Discussion

The purpose of this study was to measure the accuracy of a pseudowords task for identifying SWs. Our findings in Catalan indicate that pseudoword spelling tasks, such as PSE-CL, have good psychometric properties and can be used in formal assessments to identify SW in schools, especially in third grade. This is consistent with previous research showing that pseudoword tasks have proven reliable for assessing spelling ability in primary school settings (Daffern & Ramful, 2020; Juul & Petersen, 2017; Rothe et al., 2015; Van Vreckem & Desoete, 2018), as they target a basic skill of spelling development (i.e., phonological abilities), allowing to detect threshold scores for children who struggle in writing activities (Larkin et al., 2013; Rothe et al., 2015; Van Vreckem & Desoete, 2018). SWs failed to establish grapheme–phoneme correspondences at the same proficiency levels as their peers, tracing back to a sustained difficulty with core grapheme–phoneme mapping abilities, which seem to pose an everlasting constraint, still

tangible in the latter grades of primary education (Friend & Olson, 2008; Juul & Petersen, 2017; Van Vreckem & Desoete, 2018).

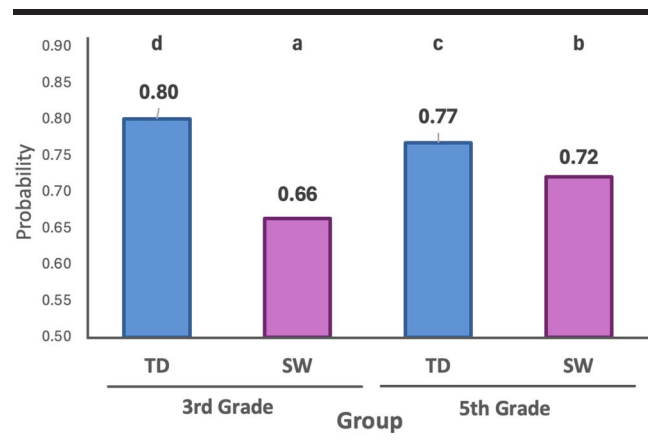
Regarding our first research question, whether a pseudoword spelling task can help distinguish SWs, our findings suggest that this task has the potential to discriminate between SWs and TD peers in the classroom. All task-reliability measures indicated fair performance, specifically a low-difficulty threshold close to ceiling effects. Nonetheless, it is not surprising that, for the general population, this task might be oversimplified and readily accessible, given their developmental trajectories.

When considering written performance on a discriminatory task, we found that the task became relevant in the early grades, such as third grade, when writing skills are still consolidating and foundational processes are more pronounced (Fitzgerald & Shanahan, 2000). At this intermediate stage, PSE-CL has demonstrated an excellent discriminatory ability, but in later grades (i.e., fifth grade), we observed a decrease in its sensitivity. This decrease can be interpreted as evidence that Catalan is an intermediate-depth language (Share, 2008), in which statistical learning plays a key role (Kemp et al., 2015; Treiman, 2017). Children from intermediate-depth languages cannot rely solely on grapheme–phoneme correspondences and need to recruit both lexical and sublexical knowledge to write (Ziegler & Goswami, 2005). Thus, they develop a sensitivity to frequent patterns and orthographic distributions that allows them to quickly recognize and retrieve probable orthographic combinations (Chetail, 2017; Treiman, 2017).

Importantly, SWs from both grades were outperformed by their TD peers, but the difference was more substantial in third grade. TD third graders ranked best on the task, possibly due to recently consolidated, explicitly taught spelling rules and the high instructional focus at this stage. As students progress, pseudoword spelling tends to shift from reliance on sublexical phoneme–grapheme mappings toward higher order patterns and stored orthographic representations, thereby reducing the diagnostic power of tasks that primarily assess phonological conversion. As seen in other intermediate or opaque-depth languages (Portuguese: Fernandes et al., 2008; French: Sprenger-Charolles et al., 1998), orthographic knowledge contributes to the construction of final orthographic representations (Fernandes et al., 2008; C. Perry et al., 2002). Building a complete orthographic representation does not rely only on converting sounds into letters in a strict way. The process is not all-or-none: Writers also use larger, meaningful units, such as letter sequences or morphemes, rather than depending solely on individual letters (C. Perry & Ziegler, 2004; Treiman et al., 2002).

Furthermore, the use of morphophonological correspondences (Nunes & Bryant, 2009) supports the development

Figure 1. Predicted mean probabilities of accurate spelling by group and grade. Latin letters for marginal mean ranking from the lowest “a” and upward. TD = typically developing; SW = struggling writers.



of sensitivity to orthographic knowledge and contextual spelling patterns, similar to those observed in opaque languages (Deacon et al., 2008; Treiman et al., 2002). Therefore, given that a pseudospelling task aims to minimize lexical support and force writers to rely on phonological-mapping spelling processes, it is not surprising that it is most effective when used as a screening measure for younger learners rather than as a primary screening measure for older students. Nonetheless, sublexical graphotactical sensibility is present from early ages (Kemp et al., 2015; Treiman, 1993, 2017), so it should be highlighted that adopting a multimeasure approach to identification is important, especially in later grades, where reliance on a single task may fail to capture the complexity and heterogeneity of writing challenges. Future research should examine how pseudospelling performance interplays with other transcription and text-level measures across development to refine screening practices for SWs.

Our second hypothesis entailed how the performance on the task could be conditioned by the item's features. Pseudowords have to be processed via the sublexical pathway (as per the dual-route model), and thus, specific manipulations of item features were expected to serve as benchmark effects to better understand the sublexical constraints. Nonetheless, we did not find an item-based characteristic (i.e., either length or frequency) that could be influencing the performance. This result does not mean that item features are irrelevant. Instead, it suggests that most of the differences in performance come from individual abilities rather than from the items themselves. When we account for group or grade differences, there is very little variation left at the item level to explain features such as length or frequency, because the main source of variation is absorbed by differences between participants. In fact, this interpretation aligns with the notion that performance on pseudowords may reflect an individual-level skill rather than item-specific properties. Those who have generalized orthographic and phonological rules tend to apply them successfully across items, regardless of their specific characteristics, whereas individuals who have not automatized these processes may struggle consistently (Mlakar et al., 2025). This is particularly relevant because pseudowords have traditionally been used to assess the degree of automatization of spelling processes (Daffern & Ramful, 2020; Martínez-Tomás et al., 2025).

In this case, our exploratory study cannot disentangle the item's features at play, but we have seen indices of superficial differences in performance (see Table A2 of the Appendix for a breakdown of the performance by items). Several studies (e.g., with children with dyslexia) have pointed out that length effects may be due to the syllabic complexity of longer stimuli (Juul & Petersen, 2017; Martens & de Jong, 2006), and thus, future tasks should aim to test these effects

in more pronounced differences of length. Complementarily, the frequency of the base word was expected to play a role in the spelling decisions, due to the influence of pre-established morphophonological correspondences (Nunes & Bryant, 2009) as an alternate access to spelling, indicating reliance on contextualized spelling patterns (Deacon et al., 2008; Treiman et al., 2002). Future research should expand our results and explore the performance of this task in other grades and contexts.

The Triple Word Form Theory (Bahr, 2015) proposes that spelling and reading involve the interaction of three types of linguistic representations: phonological (sounds), orthographic (letters and written patterns), and morphological (meaningful units such as roots and affixes). This means that sublexical spelling decisions would not rely solely on phoneme-grapheme mapping strategies; they can also draw on phonological, morphological, and orthographic knowledge, and these sources may interact dynamically (Bitan et al., 2007; Sprenger-Charolles et al., 2003). Future studies, thus, should explore how and when these dynamics of the different knowledge interact and how they impact the developmental trajectories for SWs. In this sense, we consider that the orthographic depth (the degree of transparency between sounds and letters) could also play a role in how these forms are accessed (Tolchinsky et al., 2012; Ziegler & Goswami, 2005). Sublexical correspondences and language regularities may favor rapid, undifferentiated performance early on in transparent languages, whereas semitransparent and opaque languages will require the mobilization of early lexical encoding skills (Caravolas, 2004; Fernandes et al., 2008; Sprenger-Charolles et al., 2006).

Practical Implications

Although a substantial body of research has focused on identifying reading difficulties, writing skills (and particularly spelling) have received far less attention. We propose that this task can serve as a practical classroom screening tool, thanks to its straightforward and time-efficient administration and scoring. Its simplicity allows teachers or classroom assistants to implement it in a natural learning context, minimizing disruption and reducing the need for referrals to external professionals. Dictation-based tasks offer an ecologically valid approach that closely reflects everyday classroom activities and facilitates seamless integration into educational practice. This not only preserves instructional time but also embeds assessment within authentic learning experiences.

Moreover, the task aligns with the principles of Universal Design for Learning by providing flexible and accessible means of action and expression (Rose et al., 2018). Using pseudowords enables students to demonstrate phoneme-to-grapheme conversion skills without relying on lexical memory, thereby reducing cognitive and

linguistic barriers and promoting inclusivity for learners with diverse backgrounds and abilities. The format can be adapted in terms of modality (e.g., handwriting, typing, or speech to text), pacing, and visual support, ensuring participation for students with motor, auditory, or processing challenges. By incorporating these features, the assessment fosters equitable opportunities to demonstrate spelling competence within authentic contexts.

Importantly, spelling involves concepts and skills that teachers routinely address in their daily practice, making the interpretation of results straightforward and actionable. Thus, the task is relevant not only for students but also for professionals from diverse educational backgrounds, who can easily extract meaningful insights. Nonetheless, it has to be applied with caution, as it does not substitute a full battery of diagnostic tasks, and it should be used as a part of formal assessment or a mere orientation to detect the need for further literacy support, either (a) reinforcing phonographemic activities in class, (b) following response to intervention methods (Gillon et al., 2023), or (c) referring students to a speech-language pathologist.

Limitations

Our current study is not without limitations. Results should be interpreted with caution due to the particular linguistic setting of Catalan. Further studies should explore other grades and performance across other school levels, to establish more accurate developmental and longitudinal markers. Moreover, larger samples that account for SES differences should also be considered. A broader representation of socioeconomic strata is more representative of societies, as previous research has shown that school and parental SES (L. B. Perry et al., 2022; Salas & Pascual, 2023) impact school achievement. Moreover, it should be noted that this screening task serves as a tool for further follow-up but does not treat or reveal underlying conditions that might have led to poor performance. We recommend further exploration of dynamic assessment measures (Hasson et al., 2013; Lidz & Peña, 2009).

Conclusions

The PSE-CL is a useful task with good psychometric properties for identifying SWs in primary education, especially in third grade. While these findings are based on a relatively small sample, the results provide encouraging evidence of the screener's utility, and further research could expand these preliminary findings. Grapheme–phoneme mapping is an ability assumed to be mastered during early schooling, and that progressively gives way to morphological and orthographic strategies throughout schooling. Nonetheless, we found evidence that, at the later grades of

primary education, phonological correspondences may still constrain SW cohorts.

Author Contributions

Mariona Pascual: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Mario Figueroa:** Conceptualization, Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Naymé Salas:** Conceptualization, Methodology, Funding acquisition, Project administration, Resources, Validation, Writing – original draft, Writing – review & editing.

Data Availability Statement

Data contain potentially sensitive information of underaged children. Thus, due to privacy and ethical concerns, the data sets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments

This research was supported by Recercaixa-Associació Catalana d'Universitats Públiques (ACUP) program Grant 2015 ACUP 00175, Ministerio de Ciencia Innovación y Universidades Grants PID2019-108791GA-I00 and PID2022-137868NB-I00, and Agència de Gestió d'Ajuts Universitaris i de Recerca support to research groups Grant 2021-SGR-00205, awarded to Naymé Salas. Mariona Pascual is supported by a Beatriz Galindo junior fellowship (BG23/00055) from the Spanish Ministerio de Ciencia Innovación y Universidades, as well as the Universitat de les Illes Balears. The authors gratefully acknowledge the Institut de Recerca i Innovació Educativa for providing financial support that contributed to the publication of this work. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the article. The authors want to thank the research assistant's team for their work, the Rugged Group for their advice on best practices in R, and Gabriel Lieberman and Amir Hefetz for their statistical advice. Special thanks to the schools and participants for contributing to this project.

References

- Adelman, J. S., Johnson, R. L., McCormick, S. F., McKague, M., Kinoshita, S., Bowers, J. S., & Davis, C. J. (2014). A behavioral database for masked form priming. *Behavior Research Methods*, 46(4), 1052–1067. <https://doi.org/10.3758/s13428-013-0442-y>

- Alves, R. A., Limpo, T., Salas, N., & Joshi, R. M. (2018). Handwriting and spelling. In S. Graham, C. MacArthur, & M. Hebert (Eds.), *Best practices in writing instruction* (Vol. 3, pp. 211–239). Guilford Press. <https://www.guilford.com/books/Best-Practices-in-Writing-Instruction/Graham-MacArthur-Hebert/9781462537969/contents>
- Andrews, S., Veldre, A., & Clarke, I. E. (2020). Measuring lexical quality: The role of spelling ability. *Behavior Research Methods*, 52(6), 2257–2282. <https://doi.org/10.3758/s13428-020-01387-3>
- Angelelli, P., Marinelli, C. V., Putzolu, A., Notarnicola, A., Iaia, M., & Burani, C. (2018). Learning to spell in a language with transparent orthography: Distributional properties of orthography and whole-word lexical processing. *Quarterly Journal of Experimental Psychology*, 71(3), 704–716. <https://doi.org/10.1080/17470218.2016.1275715>
- Arfe, B., Dockrell, J., & Berninger, V. W. (2014). *Writing development in children with hearing loss, dyslexia, or oral language problems: Implications for assessment and instruction*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199827282.001.0001>
- Arfé, B., & Dockrell, J. E. (2023). Difficulties in acquiring and developing writing skills. In A. G. Spinillo & C. Sotomayor (Eds.), *Development of writing skills in children in diverse cultural contexts* (pp. 223–240). Springer. https://doi.org/10.1007/978-3-031-29286-6_10
- Bahr, R. H. (2015). Spelling strategies and word formation processes. In R. Bahr & E. Silliman (Eds.), *Routledge handbook of communication disorders* (pp. 193–203). Routledge. <https://doi.org/10.4324/9780203569245>
- Balade, J., Rodríguez, C., & Jiménez, J. E. (2024). Curriculum-based measurement for early writing struggles in kindergarten: A systematic review. *Psicologia Educativa*, 30(2), 101–110. <https://doi.org/10.5093/psed2024a11>
- Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (2012). *Words their way: Word study for phonics, vocabulary, and spelling instruction* (5th ed.). Pearson.
- Berninger, V. W., & Winn, W. D. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 96–114). Guilford Press.
- Berninger, V. W., Yates, C. M., Cartwright, A. C., Rutberg, J., Remy, E., & Abbott, R. D. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing*, 4(3), 257–280. <https://doi.org/10.1007/BF01027151>
- Bitan, T., Burman, D. D., Chou, T. L., Lu, D., Cone, N. E., Cao, F., Bigio, J. D., & Booth, J. R. (2007). The interaction between orthographic and phonological information in children: An fMRI study. *Human Brain Mapping*, 28(9), 880–891. <https://doi.org/10.1002/hbm.20313>
- Bonet, M. E., & Lloret, M. R. (1998). *Fonologia catalana*. Editorial Ariel.
- Bourassa, D. C., & Treiman, R. (2008). Morphological constancy in spelling: A comparison of children with dyslexia and typically developing children. *Dyslexia*, 14(3), 155–169. <https://doi.org/10.1002/dys.368>
- Cafri, G., Hedeker, D., & Aarons, G. A. (2015). An introduction and integration of cross-classified, multiple membership, and dynamic group random-effects models. *Psychological Methods*, 20(4), 407–421. <https://doi.org/10.1037/met0000043>
- Caravolas, M. (2004). Spelling development in alphabetic writing systems: A cross-linguistic perspective. *European Psychologist*, 9(1), 3–14. <https://doi.org/10.1027/1016-9040.9.1.3>
- Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*, 45(4), 751–774. <https://doi.org/10.1006/jmla.2000.2785>
- Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavsky, M., Onochie-Quintanilla, E., Salas, N., Schöffelová, M., Defior, S., Mikulajová, M., Seidlová-Málková, G., & Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, 23(6), 678–686. <https://doi.org/10.1177/0956797611434536>
- Caravolas, M., Volín, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology*, 92(2), 107–139. <https://doi.org/10.1016/j.jecp.2005.04.003>
- Carrera-Sabaté, J., Pons-Moll, C., & Solà, J. (2014). *Els sons de català* [The sounds of Catalan]. <https://www.ub.edu/sonscatala>
- Cassar, M., Treiman, R., Moats, L., Pollo, T., & Kessler, B. (2005). How do the spellings of children with dyslexia compare with those of nondyslexic children? *Reading and Writing*, 18(1), 27–49. <https://doi.org/10.1007/s11145-004-2345-x>
- Cataldo, S., & Ellis, N. (1988). Interactions in the development of spelling, reading and phonological skills. *Journal of Research in Reading*, 11(2), 86–109. <https://doi.org/10.1111/j.1467-9817.1988.tb00153.x>
- Chetail, F. (2017). What do we do with what we learn? Statistical learning of orthographic regularities impacts written word processing. *Cognition*, 163, 103–120. <https://doi.org/10.1016/j.cognition.2017.02.015>
- Choi, Y.-J., & Asilkalkan, A. (2019). R packages for item response theory analysis: Descriptions and features. *Measurement: Interdisciplinary Research and Perspectives*, 17(3), 168–175. <https://doi.org/10.1080/15366367.2019.1586404>
- Coker, D. L., Jr., Ritchey, K. D., Uribe-Zarain, X., & Jennings, A. S. (2018). An analysis of first-grade writing profiles and their relationship to compositional quality. *Journal of Learning Disabilities*, 51(4), 336–350. <https://doi.org/10.1177/0022219417708171>
- Cragg, L., & Nation, K. (2006). Exploring written narrative in children with poor reading comprehension. *Educational Psychology*, 26(1), 55–72. <https://doi.org/10.1080/01443410500340991>
- Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2010). *Bateria de Evaluación de los Procesos Lectores, Revisada* [Revised Assessment Battery of Reading Processes, PROLEC-R]. TEA. <https://web.teaediciones.com/PROLEC-R-Bateria-de-Evaluacion-de-los-Procesos-Lectores-Revisad.aspx>
- Daffern, T., Mackenzie, N. M., & Hemmings, B. (2015). The development of a spelling assessment tool informed by triple word form theory. *The Australian Journal of Language and Literacy*, 38(2), 72–82. <https://doi.org/10.1007/BF03651958>
- Daffern, T., & Ramful, A. (2020). Measurement of spelling ability: Construction and validation of a phonological, orthographic and morphological pseudo-word instrument for students in Grades 3–6. *Reading and Writing*, 33(3), 571–603. <https://doi.org/10.1007/s11145-019-09976-1>
- Deacon, S. H., Parrila, R., & Kirby, J. R. (2008). A review of the evidence on morphological processing in dyslexics and poor readers: A strength or weakness? In G. Reid, A. J. Fawcett, F. Manis, & L. S. Siegel (Eds.), *The SAGE handbook of dyslexia* (pp. 212–237). SAGE Publications. <https://doi.org/10.4135/9780857020987.n11>
- de Bree, E., & van den Boer, M. (2019). Knowing what we don't know: Cognitive correlates of early spelling of different target

- types. *Reading and Writing*, 32, 2125–2148. <https://doi.org/10.1007/s11145-019-09936-9>
- Denman, D., Cordier, R., Kim, J.-H., Munro, N., & Speyer, R. (2021). What influences speech-language pathologists' use of different types of language assessments for elementary school-age children? *Language, Speech, and Hearing Services in Schools*, 52(3), 776–793. https://doi.org/10.1044/2021_LSHSS-20-00053
- Diamanti, V., Goulandris, N., Stuart, M., & Campbell, R. (2014). Spelling of derivational and inflectional suffixes by Greek-speaking children with and without dyslexia. *Reading and Writing*, 27(2), 337–358. <https://doi.org/10.1007/s11145-013-9447-2>
- Dockrell, J., Connelly, V., Walter, K., & Critten, S. (2018). The role of curriculum based measures in assessing writing products. In B. Miller, P. McCardle, & V. Connelly (Eds.), *Writing development in struggling learners: Understanding the needs of writers across the lifecourse* (pp. 182–215). Brill Academic Publishers. https://dx.doi.org/10.1163/9789004346369_011
- Dockrell, J. E., Forrest, C. L., Law, J., Mathers, S., & Charlton, J. (2022). Screening for language difficulties in disadvantaged populations on entry to early years education: Challenges and opportunities. *Frontiers in Pediatrics*, 10, Article 833603. <https://doi.org/10.3389/fped.2022.833603>
- Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in young children. *Child and Adolescent Mental Health*, 20(2), 116–125. <https://doi.org/10.1111/camh.12072>
- Dockrell, J. E., Ricketts, J., Charman, T., & Lindsay, G. (2014). Exploring writing products in students with language impairments and autism spectrum disorders. *Learning and Instruction*, 32, 81–90. <https://doi.org/10.1016/j.learninstruc.2014.01.008>
- Dockrell, J. E., Ricketts, J., Palikara, O., Charman, T., & Lindsay, G. A. (2019). What drives educational support for children with developmental language disorder or autism spectrum disorder: Needs, or diagnostic category? *Frontiers in Education*, 4, Article 29. <https://doi.org/10.3389/feduc.2019.00029>
- Dunn, M. (2021). The challenges of struggling writers: Strategies that can help. *Education Sciences*, 11(12), Article 795. <https://doi.org/10.3390/educsci11120795>
- Fernandes, S., Ventura, P., Querido, L., & Morais, J. (2008). Reading and spelling acquisition in European Portuguese: A preliminary study. *Reading and Writing*, 21(8), 805–821. <https://doi.org/10.1007/s11145-007-9093-7>
- Firas Dib. (2021). *Regex101: Build, test, and debug regex*. <https://regex101.com/>
- Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, 35(1), 39–50. https://doi.org/10.1207/S15326985EP3501_5
- Friend, A., & Olson, R. K. (2008). Phonological spelling and reading deficits in children with spelling disabilities. *Scientific Studies of Reading*, 12(1), 90–105. <https://doi.org/10.1080/10888430701773876>
- Gagolewski, M. (2022). stringi: Fast and portable character string processing in R. *Journal of Statistical Software*, 103(2), 1–59. <https://doi.org/10.18637/jss.v103.i02>
- Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56. [https://doi.org/10.1016/0049-089X\(92\)90017-B](https://doi.org/10.1016/0049-089X(92)90017-B)
- Gil, V., de León, S. C., & Jiménez, J. E. (2021). Universal screening for writing risk in Spanish-speaking first graders. *Reading & Writing Quarterly*, 37(2), 117–135. <https://doi.org/10.1080/10573569.2020.1733451>
- Gillon, G., McNeill, B., Scott, A., Arrow, A., Gath, M., & Macfarlane, A. (2023). A better start literacy approach: Effectiveness of Tier 1 and Tier 2 support within a response to teaching framework. *Reading and Writing*, 36(3), 565–598. <https://doi.org/10.1007/s11145-022-10303-4>
- Giovagnoli, S., Marotta, L., Magri, S., Muccinelli, M., Albani, A., Casu, G., Garofalo, S., & Benassi, M. (2020). Preliminary validation of the CI-FRA checklist: A simple screening tool for measuring the early signs of reading and spelling disorders in Italian primary students. *Frontiers in Psychology*, 11, Article 516424. <https://doi.org/10.3389/fpsyg.2020.516424>
- Goswami, U. (2001). Early phonological development and the acquisition of literacy. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (Vol. 1, pp. 111–125). Guilford Press. <https://www.guilford.com/books/Handbook-of-Early-Literacy-Research-Volume-1/Neuman-Dickinson/9781572308954/contents>
- Graham, S., Harris, K., & Adkins, M. (2018). The impact of supplemental handwriting and spelling instruction with first grade students who do not acquire transcription skills as rapidly as peers: A randomized control trial. *Reading and Writing*, 31(6), 1273–1294. <https://doi.org/10.1007/s11145-018-9822-0>
- Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing*, 27(9), 1703–1743. <https://doi.org/10.1007/s11145-014-9517-0>
- Grainger, J., Spinelli, E., & Ferrand, L. (2000). Effects of baseword frequency and orthographic neighborhood size in pseudohomophone naming. *Journal of Memory and Language*, 42(1), 88–102. <https://doi.org/10.1006/jmla.1999.2675>
- Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal*, 43(2), 295–340. <https://doi.org/10.3102/00028312043002295>
- Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2013). Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Language Teaching and Therapy*, 29(1), 57–75. <https://doi.org/10.1177/0265659012459526>
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But.... *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
- Helsel, L., & Greenberg, D. (2007). Helping struggling writers succeed: A self-regulated strategy instruction program. *The Reading Teacher*, 60(8), 752–760. <https://doi.org/10.1598/RT.60.8.5>
- Hoefflin, G., & Franck, J. (2005). Development of spelling skills in children with and without learning disabilities. *Studies in Language and Literature*, 5(2), 175–192. <https://doi.org/10.1007/s10674-005-0917-6>
- Joye, N., Dockrell, J. E., & Marshall, C. R. (2020). The spelling errors of French and English children with developmental language disorder at the end of primary school. *Frontiers in Psychology*, 11, Article 1789. <https://doi.org/10.3389/fpsyg.2020.01789>
- Juul, H., & Petersen, D. K. (2017). Length effects in pseudo-word spelling: Stronger in dyslexic than in non-dyslexic students. *Annals of Dyslexia*, 67(3), 369–382. <https://doi.org/10.1007/s11881-017-0149-3>
- Kamvar, Z. N. (2020). *Matchmaker R package* [Computer program]. <https://www.repidemicsconsortium.org/matchmaker/>
- Kemp, N., Treiman, R., Blackley, H., Svoboda, I., & Kessler, B. (2015). Lexical classification and spelling: Do people use atypical spellings for atypical pseudowords? *Reading and Writing*, 28(8), 1187–1202. <https://doi.org/10.1007/s11145-015-9567-y>

- Kim, Y. S., Petscher, Y., Foorman, B. R., & Zhou, C. (2010). The contributions of phonological awareness and letter-name knowledge to letter-sound acquisition—A cross-classified multilevel model approach. *Journal of Educational Psychology*, 102(2), 313–326. <https://doi.org/10.1037/a0018449>
- Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language*, 102(1), 30–45. <https://doi.org/10.1016/j.bandl.2006.11.001>
- Larkin, R. F., & Snowling, M. J. (2008). Comparing phonological skills and spelling abilities in children with reading and language impairments. *International Journal of Language & Communication Disorders*, 43(1), 111–124. <https://doi.org/10.1080/13682820601178584>
- Larkin, R. F., Williams, G. J., & Blaggan, S. (2013). Delay or deficit? Spelling processes in children with specific language impairment. *Journal of Communication Disorders*, 46(5–6), 401–412. <https://doi.org/10.1016/j.jcomdis.2013.07.003>
- Lidz, C. S., & Peña, E. D. (2009). Response to intervention and dynamic assessment: Do we just appear to be speaking the same language? *Seminars in Speech and Language*, 30(2), 121–133. <https://doi.org/10.1055/s-0029-1215719>
- Llaurado, A., & Tolchinsky, L. (2016). The developmental pattern of spelling in Catalan from first to fifth school grade. *Writing Systems Research*, 8(1), 64–83. <https://doi.org/10.1080/17586801.2014.1000812>
- Mahfouz, M. A. S., & Mohamed, A. H. H. (2023). Using word reading fluency curriculum-based measurements to monitor students' reading progress in Grade 2. *Education Sciences*, 13(2), Article 217. <https://doi.org/10.3390/educsci13020217>
- Martens, V., & de Jong, P. F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and Language*, 98(2), 140–149. <https://doi.org/10.1016/j.bandl.2006.04.003>
- Martin, K. I., Lawson, E., Carpenter, K., & Hummer, E. (2020). English word and pseudoword spellings and phonological awareness: Detailed comparisons from three L1 writing systems. *Frontiers in Psychology*, 11, Article 1309. <https://doi.org/10.3389/fpsyg.2020.01309>
- Martínez-Tomás, C., Bacierno, A., Lázaro, M., & Hinojosa, J. A. (2025). What do pseudowords tell us about word processing? An overview. *Frontiers in Language Sciences*, 4, Article 1504770. <https://doi.org/10.3389/flang.2025.1504770>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Erlbaum.
- McLeod, S., & McCormack, J. (2015). What are speech, language and literacy? In *Introduction to speech, language and literacy* (pp. 1–39). Oxford University Press.
- Meade, G., Grainger, J., Midgley, K. J., Emmorey, K., & Holcomb, P. J. (2018). From sublexical facilitation to lexical competition: ERP effects of masked neighbor priming. *Brain Research*, 1685, 29–41. <https://doi.org/10.1016/j.brainres.2018.01.029>
- Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, 138(2), 322–352. <https://doi.org/10.1037/a0026744>
- Mrakar, H., Hirst-Plein, J., & Koch, M. J. (2025). Cognitive and linguistic influences on EFL real word and pseudoword spelling: Predictors and error analysis. *Languages*, 10(5), Article 93. <https://doi.org/10.3390/languages10050093>
- Muthén, L. K., & Muthén, B. O. (2025). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Nahm, F. S. (2022). Receiver operating characteristic curve: Overview and practical use for clinicians. *Korean Journal of Anesthesiology*, 75(1), 25–36. <https://doi.org/10.4097/kja.21209>
- Nedungadi, P., Shah, S. M., Stokes, M. A., Kumar Nair, V., Moorkoth, A., & Raman, R. (2024). Mapping autism's research landscape: Trends in autism screening and its alignment with sustainable development goals. *Frontiers in Psychiatry*, 14, Article 1294254. <https://doi.org/10.3389/fpsy.2023.1294254>
- Nielsen, A.-M. V., & Juul, H. (2016). Predictors of early versus later spelling development in Danish. *Reading and Writing*, 29(2), 245–266. <https://doi.org/10.1007/s11145-015-9591-y>
- Nunes, T., & Bryant, P. (2009). *Children's reading and spelling: Beyond the first steps*. Wiley.
- Ozubko, J. D., & Joordens, S. (2011). The similarities (and familiarities) of pseudowords and extremely high-frequency words: Examining a familiarity-based explanation of the pseudoword effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 123–139. <https://doi.org/10.1037/a0021099>
- Pascual, M., Naymé, S., Ribas Seix, T., & Cros Alavedra, A. (2020). *Children at-risk of writing difficulties* [Doctoral dissertation, Autonomous University of Barcelona]. <https://ddd.uab.cat/record/240935>
- Pascual, M., & Salas, N. (2021). The cognitive profile and text-based traits of struggling writers (*El perfil cognitivo y los rasgos textuales de los escritores con dificultades*). *Journal for the Study of Education and Development: Infancia y Aprendizaje*, 44(1), 219–253. <https://doi.org/10.1080/02103702.2020.1848090>
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–214). John Benjamins. <https://doi.org/10.1075/swll.11.14per>
- Perry, C., & Ziegler, J. C. (2004). Beyond the two-strategy model of skilled spelling: Effects of consistency, grain size, and orthographic redundancy. *The Quarterly Journal of Experimental Psychology*, 57(2), 325–356. <https://doi.org/10.1080/02724980343000323>
- Perry, C., Ziegler, J. C., & Coltheart, M. (2002). How predictable is spelling? Developing and testing metrics of phoneme-grapheme contingency. *The Quarterly Journal of Experimental Psychology*, 55(3), 897–915. <https://doi.org/10.1080/02724980143000640>
- Perry, L. B., Saatcioglu, A., & Mickelson, R. A. (2022). Does school SES matter less for high-performing students than for their lower-performing peers? A quantile regression analysis of PISA 2018 Australia. *Frontiers in Education*, 10(1), Article 17. <https://doi.org/10.1186/s40536-022-00137-5>
- Piercy, W., & Dockrell, J. E. (2023). Assessing children's writing products using curriculum-based measures of writing (CBM-W). *Educational & Child Psychology*, 40(1), 44–60. <https://doi.org/10.53841/bpsecp.2023.40.1.44>
- Protopapas, A., Aikaterini, F., Styliani, D., Skaloumbakas, C., & Angeliki, M. (2013). What do spelling errors tell us? Classification and analysis of errors made by Greek schoolchildren with and without dyslexia. *Reading and Writing*, 26(5), 615–646. <https://doi.org/10.1007/s11145-012-9378-3>
- Quick, N. A., Harrison, M., & Erickson, K. (2021). A multilingual spelling analysis of children who are hard of hearing. *Journal of Deaf Studies and Deaf Education*, 26(1), 112–129. <https://doi.org/10.1093/deafed/enaa021>
- Rafel, J. (1996). *Dictionary of Frequencies*. Institut d'Estudis Catalans. <https://latel.upf.edu/terminotica/diccfreq.htm>

- R Core Team.** (2020). *R: A language and environment for statistical computing* (Version 3.6.3) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Revelle, W., & Condon, D. M.** (2019). Reliability from α to ω : A tutorial. *Psychological Assessment*, 31(12), 1395–1411. <https://doi.org/10.1037/pas0000754>
- Ritchey, K. D., & Coker, D. L., Jr.** (2014). Identifying writing difficulties in first grade: An investigation of writing and reading measures. *Learning Disabilities Research & Practice*, 29(2), 54–65. <https://doi.org/10.1111/ldrp.12030>
- Robitzsch, A., Kiefer, T., & Wu, M.** (2024). *TAM: Test analysis modules* (R Package, Version 4.3-25). <https://CRAN.R-project.org/package=TAM>
- Rose, D. H., Robinson, K. H., Hall, T. E., Coyne, P., Jackson, R. M., Stahl, W. M., & Wilcauskas, S. L.** (2018). Accurate and informative for all: Universal Design for Learning (UDL) and the future of assessment. In S. Elliott, R. Kettler, P. Beddow, & A. Kurz (Eds.), *Handbook of accessible instruction and testing practices* (pp. 167–180). Springer. https://doi.org/10.1007/978-3-319-71126-3_11
- Rothe, J., Cornell, S., Ise, E., & Schulte-Körne, G.** (2015). A comparison of orthographic processing in children with and without reading and spelling disorder in a regular orthography. *Reading and Writing*, 28(9), 1307–1332. <https://doi.org/10.1007/s11145-015-9572-1>
- Sacre, L., & Masterson, J.** (2000). *Single Word Spelling Test (SWST)*. NFER-Nelson.
- Salas, N.** (2020). Non-phonological strategies in spelling development. *Frontiers in Psychology*, 11, Article 1071. <https://doi.org/10.3389/fpsyg.2020.01071>
- Salas, N., & Pascual, M.** (2023). Impact of school SES on literacy development. *PLOS ONE*, 18(12), Article e0295606. <https://doi.org/10.1371/journal.pone.0295606>
- Seidenberg, M. S., Farry-Thorn, M., & Zevin, J. D.** (2022). Models of word reading. In M. J. Snowling, C. Hulme, & K. Nation (Eds.), *The science of reading: A handbook* (2nd ed., pp. 36–59). Wiley. <https://doi.org/10.1002/9781119705116.ch2>
- Share, D. L.** (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin*, 134(4), 584–615. <https://doi.org/10.1037/0033-2909.134.4.584>
- Silliman, E. R., Bahr, R. H., Nagy, W., & Berninger, V.** (2018). Language bases of spelling in writing during early and middle childhood: Grounding applications to struggling writers in typical writing development. In B. Miller, P. McCardle, & V. Connelly (Eds.), *Writing development in struggling learners* (pp. 99–119). Brill. https://doi.org/10.1163/9789004346369_007
- Snyder, S. M., & Ayres, K.** (2020). Investigating the usage of reading curriculum-based measurement (CBM-R) to formatively assess the basic reading skills of students with intellectual disability. *Education and Training in Autism and Developmental Disabilities*, 55(1), 60–74. <https://doi.org/10.1177/215416472005500106>
- Solà, J., Lloret, M. R., Mascaró, J., & Pérez Saldanya, M.** (2008). *Gramàtica del català contemporani* [Contemporary Catalan grammar] (Vol. 1). Empúries.
- Sprenger-Charolles, L., Colé, P., & Serniclaes, W.** (2006). *Reading acquisition and developmental dyslexia*. Psychology Press. <https://doi.org/10.4324/9780203759714>
- Sprenger-Charolles, L., Siegel, L. S., Béchennec, D., & Serniclaes, W.** (2003). Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology*, 84(3), 194–217. [https://doi.org/10.1016/S0022-0965\(03\)00024-9](https://doi.org/10.1016/S0022-0965(03)00024-9)
- Sprenger-Charolles, L., Siegel, L. S., & Bonnet, B.** (1998). Reading and spelling acquisition in French: The role of phonological mediation and orthographic factors. *Journal of Experimental Child Psychology*, 68(2), 134–165. <https://doi.org/10.1006/jecp.1997.2422>
- Suárez-Coalla, P., Villanueva, N., González-Pumariega, S., & González-Nosti, M.** (2016). Spelling difficulties in Spanish-speaking children with dyslexia/Dificultades de escritura en niños españoles con dislexia. *Journal for the Study of Education and Development: Infancia y Aprendizaje*, 39(2), 275–311. <https://doi.org/10.1080/02103702.2015.1132979>
- Ten Berge, J. M. F., & Sočan, G.** (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, 69(4), 613–625. <https://doi.org/10.1007/BF02289858>
- Thiele, C., & Hirschfeld, G.** (2021). cutpointr: Improved estimation and validation of optimal cutpoints in R. *Journal of Statistical Software*, 98(11), 1–27. <https://doi.org/10.18637/jss.v098.i11>
- Tolchinsky, L.** (2021). Linguistic patterns of spelling of isolated words to dictation and text-composing in Catalan across elementary school (*Patrones lingüísticos de la ortografía en el dictado de palabras aisladas y en la composición de textos en catalán en la escuela primaria*). *Journal for the Study of Education and Development: Infancia y Aprendizaje*, 44(1), 183–218. <https://doi.org/10.1080/02103702.2020.1848091>
- Tolchinsky, L., Ribera, P., & García-Parejo, I.** (2012). Conocimientos iniciales y logros de los alumnos en el aprendizaje de la lectura y la escritura [Initial knowledge and pupils' achievement in literacy learning]. *Cultura y Educación*, 24(4), 415–433. <https://doi.org/10.1174/113564012803998811>
- Treiman, R.** (1993). *Beginning to spell: A study of first-grade children*. Oxford University Press. <https://doi.org/10.1093/oso/9780195062199.001.0001>
- Treiman, R.** (2017). Learning to spell: Phonology and beyond. *Cognitive Neuropsychology*, 34(3–4), 83–93. <https://doi.org/10.1080/02643294.2017.1337630>
- Treiman, R.** (2018). Teaching and learning spelling. *Child Development Perspectives*, 12(4), 235–239. <https://doi.org/10.1111/cdep.12292>
- Treiman, R., Kessler, B., & Bick, S.** (2002). Context sensitivity in the spelling of English vowels. *Journal of Memory and Language*, 47(3), 448–468. [https://doi.org/10.1016/S0749-596X\(02\)00010-4](https://doi.org/10.1016/S0749-596X(02)00010-4)
- Treiman, R., Kessler, B., & Caravolas, M.** (2019). What methods of scoring young children's spelling best predict later spelling performance? *Journal of Research in Reading*, 42(1), 80–96. <https://doi.org/10.1111/1467-9817.12241>
- Troia, G. A.** (2025). Writing in students with diverse abilities. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 407–425). Guilford Press. <https://www.guilford.com/books/Handbook-of-Writing-Research/MacArthur-Graham-Fitzgerald/9781462557271>
- Vale, A. P., & Perpétua, R.** (2020). Early context-conditioned orthographic knowledge in European Portuguese: The spelling of the schwa. *Frontiers in Education*, 5, Article 513577. <https://doi.org/10.3389/educ.2020.513577>
- Van den Noortgate, W., De Boeck, P., & Meulders, M.** (2003). Cross-classification multilevel logistic models in psychometrics. *Journal of Educational and Behavioral Statistics*, 28(4), 369–386. <https://doi.org/10.3102/10769986028004369>
- Van Vreckem, C., & Desoete, A.** (2018). Spelling of pseudowords and real words in Dutch-speaking children with and without dyslexia. *Topics in Language Disorders*, 38(4), 286–298. <https://doi.org/10.1097/TLD.0000000000000164>
- Verhagen, W. T., Aarnoutse, C. A. J., & van Leeuwe, J. F. J.** (2010). Spelling and word recognition in Grades 1 and 2: Relations to phonological awareness and naming speed in Dutch children. *Applied Psycholinguistics*, 31(1), 59–80. <https://doi.org/10.1017/S0142716409990166>

-
- Wheeler, M. W.** (2005). *The phonology of Catalan*. Oxford University Press. <https://doi.org/10.1093/oso/9780199258147.001.0001>
- Wickham, H., Çetinkaya-Rundel, M., & Golemund, G.** (2023). *R for data science*. O'Reilly Media.
- Wiley, R. W., Singh, S., Baig, Y., Key, K., & Purcell, J. J.** (2024). The English sublexical toolkit: Methods for indexing sound–spelling consistency. *Behavior Research Methods*, 56(7), 6826–6861. <https://doi.org/10.3758/s13428-024-02395-3>
- Williams, G. J., & Larkin, R. F.** (2013). Narrative writing, reading and cognitive processes in middle childhood: What are the links? *Learning and Individual Differences*, 28, 142–150. <https://doi.org/10.1016/j.lindif.2012.08.003>
- Wu, S., Crespi, C. M., & Wong, W. K.** (2012). Comparison of methods for estimating the intraclass correlation coefficient for binary responses in cancer prevention cluster randomized trials. *Contemporary Clinical Trials*, 33(5), 869–880. <https://doi.org/10.1016/j.cct.2012.05.004>
- Youden, W. J.** (1950). Index for rating diagnostic tests. *Cancer*, 3(1), 32–35. [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3)
- Ziegler, J. C., & Goswami, U.** (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131(1), 3–29. <https://doi.org/10.1037/0033-2909.131.1.3>

Appendix

Supplementary Sample Characterization and Detailed Accuracy Results in PSE-CL

Table A1. Mean (standard deviation) results in the identification measures for struggling writers group.

Measure	Struggling writers		Typically developing controls	
	3rd	5th	3rd	5th
Handwriting Identification Score ^a	3.31 (1.41)	6.00 (2.45)	5.93 (1.98)	8.69 (2.47)
Spelling Identification Score ^b	16.92 (10.92)	20.45 (10.68)	25.32 (3.67)	28.75 (4.05)
Text Generation Identification Score ^c	27.90 (20.59)	58.05 (32.36)	49.96 (35.38)	83.86 (39.05)

Note. All tasks were carried out 18 months prior.

^aHandwriting task measured with the alphabet task. Total score includes the number of letters written in 15" (Berninger et al., 1992). ^bSpelling task from Salas (2020) scored in total number of well-written words (34 maximum). ^cAverage of total amount of written words in two curriculum-based measurement texts (opinion essay and narrative text). Means have been displayed for complementary information. Further details can be found in Pascual and Salas (2021).

Table A2. Mean accuracy (in %) in the Pseudoword Spelling Evaluation in Catalan Language task (mean and standard deviation) per item features.

Grade		Struggling writers		Typically developing controls	
		Short psw.	Long psw.	Short psw.	Long psw.
3rd grade	High frequency	73.63 (2.31)	66.48 (3.51)	91.84 (1.38)	89.29 (2.21)
	Low frequency	70.63 (2.70)	62.50 (2.75)	91.23 (1.61)	81.84 (2.11)
5th grade	High frequency	77.60 (2.38)	74.68 (3.52)	90.18 (1.41)	80.36 (2.66)
	Low frequency	76.86 (2.72)	64.77 (2.95)	85.80 (1.86)	82.55 (1.94)

Note. psw. = pseudoword.