jove

Advancing Dyslexia Assessment in Children through Computerized Testing

Juan E. Jiménez[1], **Eduardo García**[1], **Jennifer Balade**[1]

[1] The University Institute of Neuroscience (IUNE), Universidad de La Laguna

Corresponding Author

Juan E. Jiménez

ejimenez@ull.edu.es

Abstract

The acquisition of reading skills is an intricate process that demands the cultivation of various domain-general and domain-specific abilities. Consequently, it is unsurprising that many children grapple with maintaining proficiency at the grade level, particularly when confronted with challenges spanning multiple abilities across both domains, as observed in individuals with reading difficulties. Strikingly, despite reading difficulties being among the most prevalent neurodevelopmental disorders affecting school-aged children, the majority of available diagnostic tools lack a comprehensive framework for assessing the full spectrum of cognitive skills linked to dyslexia, with minimal computerized options. Notably, there are currently limited tools with these features available for Spanish-speaking children. The aim of this study was to delineate the protocol for diagnosing Spanish-speaking children with reading difficulties using the Sicole-R multimedia battery. This tool for elementary grades focuses on evaluating cognitive skills that are associated with dyslexia as prescribed by the scientific literature. Specifically, it concentrates on assessing a range of cognitive abilities that studies have demonstrated to be linked to dyslexia. This focus is based on the observation that individuals with dyslexia typically exhibit deficits in several of the cognitive areas evaluated by this digital tool. The robust internal consistency and multidimensional internal structure of the battery were demonstrated. This multimedia battery has proven to be a fitting tool for diagnosing children with reading difficulties in primary education, offering a comprehensive cognitive profile that is valuable not only for diagnostic purposes but also for tailoring individualized instructional plans.

Introduction

Dyslexia is a neurodevelopmental disorder characterized by difficulties in accurate and/or fluent word recognition and poor spelling and decoding abilities and is characterized by unexpected and persistent difficulty in acquiring efficient reading skills despite conventional instruction, adequate intelligence, and sociocultural opportunity[1]. This

neurobiological disorder often manifests as challenges in reading, spelling, and writing, primarily due to phonological deficits[2,3]. "The importance of early identification of dyslexia cannot be overstated, as it allows for timely intervention and support[4,5]. When a student does not progress beyond Tier-3 in a response-to-intervention model, it becomes essential to conduct a more comprehensive assessment of both domain-general and domain-specific abilities associated with dyslexia, as highlighted by the scientific literature. The development of the technique presented here is grounded in the necessity of conducting thorough evaluations to ensure that appropriate interventions and support are provided. Moreover, previous studies underscore the utility of technology-based screening tools, such as web applications and computer games, in facilitating effective screening processes[6,7]. These studies collectively highlight the multifaceted nature of dyslexia, emphasizing the need for comprehensive assessment and intervention strategies to address the diverse cognitive profiles of individuals with dyslexia. Despite the prevalence of dyslexia among school-aged children, most available diagnostic tools lack a framework that comprehensively assesses both domain-general and domain-specific skills. Moreover, there are minimal computerized options, particularly for Spanish-speaking populations. This multimedia battery addresses these gaps by leveraging technology to facilitate a detailed assessment of cognitive skills linked to dyslexia.

Theoretical perspectives and cognitive deficits in Dyslexia

Various theoretical models, including phonological, rapid auditory processing, visual, magnocellular, and cerebellar theories, aim to explain the causes of dyslexia and inform interventions (see for a review)[8]. The phonological theory attributes dyslexia to difficulties in processing language sounds[9], while the rapid auditory processing theory links dyslexia to deficits in perceiving rapidly changing sounds[10]. Visual theory highlights the visual aspects of reading difficulties, and magnocellular theory points to impairments in visual and auditory processing pathways[11]. The cerebellar theory suggests that dyslexia arises from cerebellar impairments affecting motor control and cognitive functions[12]. Nicolson and Fawcett's Delayed Neural Commitment (DNC) framework posits that slower skill acquisition and delayed neural network development are central to dyslexia. Recent models, such as the multiple deficit model, propose that dyslexia is a complex disorder influenced by genetic, cognitive, and environmental factors[13,14,15]. For instance, Ring and Black[14] support the multiple deficit model, showing that both phonological and cognitive processing deficits contribute to the heterogeneity of dyslexia. Soriano-Ferrer et al.[15] conducted a study with Spanish-speaking children with developmental dyslexia (DD) and found significant impairments in naming speed, verbal working memory, and phonological awareness (PA). Similarly, Zygouris et al.[16] and Rauschenberger et al.[6] underscore the importance of cognitive screening tools in identifying these deficits, with dyslexic individuals consistently scoring lower than typically achieving peers.

Examining technological approaches in Dyslexia screening: Insights from research studies

Research on dyslexia screening has evolved with three main approaches: early detection strategies, multifaceted screening methods combining various assessments, and integrating technology for enhanced efficiency[17]. Politi-Georgousi's[18] recent systematic review highlights a shift toward more applications for intervening in dyslexia symptoms rather than screening processes, aligning with technology integration to improve reading skills

in dyslexic students. Various tools exist, such as the Dyslexia Early Screening Test (DEST) by Fawcett and Nicolson, which assesses speed, phonological skills, motor skills, cerebellar function, and knowledge[19]. "Computer-based tools have advanced, including a web application assessing reading and cognitive skills in Greek children[20] and tools by Hautala et al.[21] and Rauschenberg et al.[6] that use gaming and machine learning for early identification of dyslexia. Ahmad et al. integrated gaming with neural networks, achieving 95% accuracy in detection[22]. Studies across different orthographies underscore the importance of phonological awareness and rapid automatized naming in dyslexia identification[23,24].

Insights into Dyslexia among Spanish-speaking children

The study of dyslexia in Spanish-speaking children has been significantly advanced through the use of Sicole-R technology. Jiménez et al. demonstrated its effectiveness in assessing dyslexia across age groups, particularly in distinguishing between dyslexic and typically achieving readers based on phonological and syntactic processing during early elementary years[25]. Guzmán et al. investigated naming speed deficits in dyslexic children with phonological challenges, highlighting interactions between dyslexia and naming speed measured through tasks such as letter-RAN and number-RAN[26]. Further studies by Jiménez et al. explored phonological awareness deficits across different syllable structures[27], while Ortiz et al. investigated speech perception deficits among Spanish children with dyslexia, revealing impairments in speech perception development regardless of phonetic contrast or linguistic unit[28,29]. Jiménez et al. investigated the double-deficit hypothesis of dyslexia[30], followed by analyses of cognitive processes and gender-related disparities in dyslexia prevalence[31,32]. Rodrigo et al. explored lexical access among Spanish dyslexic children[33], and Jiménez et al. scrutinized syntactic processing deficits[34]. Finally, Jiménez et al. studied phonological and orthographical processes in dyslexic subtypes, highlighting differences in orthographic route efficiency[35]. These studies collectively enhance our understanding of the cognitive and linguistic challenges of dyslexia in Spanish-speaking populations.

The conducted studies share several common characteristics in terms of the age and background of the participating children. The children included in these studies ranged in age from 7 to 14 years. Most studies focused on primary school children aged between 7 and 12 years, except those that included children up to 14 years old, providing a sample that spans from early school years to preadolescence[31,32]. The participating children were primarily from the Canary Islands in Spain. Additionally, some studies included samples from other regions of Spain and Guatemala[31,32]. Participants were recruited from both public and private schools whose backgrounds included urban and suburban areas. The socioeconomic levels represented in these studies range from low-middle to working and middle class.

Together, these inquiries significantly advance our understanding of dyslexia's complexities, contributing to the field of dyslexia research. Adapted for use across multiple Ibero-American countries, including Spain, Guatemala, Chile, and Mexico, the tool facilitates the assessment of diagnostic accuracy and precision in a diverse Spanish-speaking sample for this study.

This study aimed to delineate a protocol for diagnosing Spanish-speaking children with reading difficulties using a specialized multimedia battery. The primary goal is to provide a comprehensive assessment tool that evaluates

both domain-general and domain-specific skills associated with dyslexia.

Experimental setup overview

The SICOLE-R was programmed in the Java 2 Platform Standard Edition (J2SE). The HSQL database engine is used as a database. The software includes 6 main modules to be evaluated: 1) perceptual processing, which includes the tasks of voicing, placing, and manner of articulation; 2) phonological processing, which includes phoneme isolation, phoneme deletion, phoneme segmentation, and phoneme blending tasks; 3) naming speed, which includes the tasks of naming speed in numbers, letters, colors and pictures; 4) orthographic processing, which includes tasks of morphological comprehension of lexemes and suffixes and homophone comprehension; 5) syntactic processing, including gender, number, function words, and grammatical structure tasks; and 6) semantic processing, which influences reading comprehension tasks through informative and narrative text. Instructions for each task, accompanied by one or two trials (depending on the task) and a demonstration, are delivered by a pedagogical agent prior to the initiation of the testing phase. The application protocol for each task is illustrated here.

Prior to administering the multimedia battery to the study sample, adaptations were made to the Spanish language modality for each country (i.e., Mexico, Guatemala, Ecuador, and Chile), including adjustments to vocabulary, images, and other relevant content. The administration conditions were the same across all Latin American countries. The administration environment had to be quiet within the school and free from noise, distractions, and interruptions. The duration of the multimedia battery administration ranged from 3-4 sessions of 30 min each, depending on the student's ability and age. Due to its database compatibility with most spreadsheet and statistical data processing systems, the evaluator can analyze the results of each child and each task. Concerning the

data collection, two distinct task types were employed: 1) tasks where the examiner records students' oral performance, noting successes and errors using an external mouse, and 2) tasks requiring students to independently select options by clicking on them.

Protocol

This protocol was conducted in accordance with the guidelines provided by the Comité de Ética de la Investigación y Bienestar Animal (Research Ethics and Animal Welfare Committee, CEIBA) at Universidad de La Laguna (ULL). The data were collected at different times according to the curriculum of each country, capturing information exclusively from students whose educational administrations, schools, and parents provided consent. The test battery used in this study is registered as intellectual property and can be accessed through a transfer agreement with the ULL. For more information on how to obtain the test battery, interested parties can contact the Office of Knowledge Transfer (OTRI) at ULL.

1. SICOLE-R installation and preparation

1. Use the following inclusion criterion for applying this tool: students from second to sixth grade. Use the following exclusion criterion: do not include students with special educational needs, referring to those who needed support and targeted educational attention due to sensory impairment or acquired neurological problems, among other factors.
NOTE: The assessment is conducted with the students individually in a quiet space with access to a computer and good lighting. Headphones with an adapted microphone will be necessary to facilitate the reception of task instructions and increase student performance.

2. To install the software on a computer, execute the file as the administrator and click on the icon to open the tool. Upon opening the application, an initial interface appears, displaying various operational options.

3. Fill in the students' information who will perform the tasks before starting the evaluation. Once their data is entered, record them in a list of students, and their information can be modified later.

4. Each time a new student is added, a test screen will appear before starting the tasks. On this screen, ask the student to tap on the boot that will appear on the screen.
NOTE: This mini-game's objective is to establish a baseline for assessing students' motor control. It allows for the monitoring of individual differences in motor speed, as the software records responses by pressing keys on the keyboard. This mini game appears only once.

5. Before beginning the assessment, ask the examiner to give the following instructions: Now, you will enter a virtual circus environment with different gates. Each gate leads to fun activities where you will participate. A clown will guide you through these gates and explain what to do in each activity. Please sit next to the examiner and listen carefully to the instructions. Make sure to do your best in each task. Enjoy exploring and completing these activities!

6. Begin the tasks by selecting the student scheduled for assessment and clicking on the **Start** section. The main menu includes 5 colored doors, each corresponding to one of the modules to be evaluated.

7. Each task starts with an introductory instruction, ask the students to pay attention to this. The students are informed about the nature of the game and instructed on how to play it, facilitated by a pedagogical agent.

Ensure the agent delivers verbal instructions for each task, exemplifying the procedure through a model and presenting examples for students to emulate.

8. Following the completion of the examples, ensure the pedagogical agent provides corresponding feedback, allowing the students to redo the example if it was executed incorrectly, and the evaluation begins.
NOTE: The pedagogical agent is a virtual character designed to guide the student through the tasks. This role requires the agent to have clear and understandable speech and the ability to provide instructional guidance, task modeling, and feedback based on student responses. The pedagogical agent must be programmed to respond consistently and accurately according to the student's inputs during the tasks.

2. SICOLE-R tasks

1. Yellow Door: Perceptual processing module using a speech perception task

   1. Open the program on the computer. Navigate to the yellow door within the program interface. Click on **One of the Available Subtasks** to initiate the task: (1) voicing, (2) manner of articulation, and (3) placing.

   2. Allow the pedagogical agent to provide instructional guidance and task modeling to the student. The agent should say: Now we will hear pairs of syllables. If the syllables are identical, press the blue button; if they are different, press the red button. Watch how I do it. Now it is your turn with these examples.

   3. Present two item examples to the student:/ba/-/pa/;/ ja/-/ka/. Instruct the students again to select the blue

circle if the pairs are identical and the red button if they are distinct.

4. Once the task is completed, ensure that the pedagogical agent provides feedback based on the student's answers in the examples. Once the examples are complete, allow the agent to instruct the student to initiate the task.

5. The procedure of selecting circles (blue or red) will remain consistent throughout the task, recording hits and misses. If the student needs to hear the pair of syllables again, click on the **Speaker Icon** to replay it once more. Note that only one additional playback per item is allowed.

2. Pink Door: Phonological awareness module

1. Phoneme segmentation task

   1. Open the program on the computer. Navigate to the pink door within the program interface. Choose **Phoneme Segmentation** to select the task.

   2. Allow the pedagogical agent to provide instructional guidance to the student. The agent should say: Now we will focus on phoneme segmentation. Watch as I demonstrate. Then, you will see two example items.

   3. Present two example items to the student. Ask the student to express their answers aloud for each word.

   4. Use the external mouse to click on the blue button for correct responses and the red button for incorrect responses. Ensure the pedagogical agent offers feedback based on the student's responses to the examples.

5. After the examples, ensure the pedagogical agent instructs the student to commence the task.

2. Phoneme blending task

   1. Enter the program on the computer. Navigate to the pink door within the program interface. Choose **Phoneme Blending** to select the task.

   2. Allow the pedagogical agent to provide instructional guidance to the student. The agent says: Now we will focus on phoneme synthesis. Watch as I demonstrate. Then, you will see two example items.

   3. Present two example items to the student. Ask the student to articulate their answers aloud for each word.

   4. Utilize the external mouse to click the blue button for accurate responses and the red button for inaccurate responses. Ensure the pedagogical agent provides feedback based on the student's responses to the examples.

   5. After the examples, allow the pedagogical agent to direct the student to initiate the task. The procedure remains consistent throughout task execution, recording both accurate and inaccurate responses.

   6. If the student needs to listen to a word again, click on the **Parrot Icon** for a replay. Note that only one replay per item is allowed.

3. Phoneme isolation task

   1. Open the program on the computer. Navigate to the pink door within the program interface. Click on **Phoneme Isolation** to select the task.

2. Allow the pedagogical agent to provide instructional guidance and task modeling to the student. The agent should say: Now we will focus on isolating phonemes. Watch as I demonstrate. Then, you will see two example items. Your task is to click on the images whose name starts with the same sound as the target word.

3. Present two example items to the student. Observe whether the student clicks on the images that begin with the same sound as the target word.

4. Ensure the pedagogical agent offers feedback based on the student's responses to the examples. After the examples, ensure the agent guides the student to commence the task.

5. The procedure remains uniform throughout the task, documenting both hits and misses. If the student requires additional listening for a word, click on the **speaker Icon** to replay it. Note that only one replay per item is allowed.

4. Phoneme deletion task

1. Open the program on the computer. Navigate to the pink door within the program interface. Select **Phoneme Deletion** to choose the task.

2. Allow the pedagogical agent to provide instructional guidance and demonstrate the task to the student. The agent should say: Now we will focus on phoneme deletion. Watch as I demonstrate. Then, you will see two example items.

3. Present two example items to the student. Ask the student to say their answer aloud for each word.

4. Click with the external mouse on the blue button for correct answers and on the red button for incorrect answers. Ensure the pedagogical agent provides feedback based on the student's responses to the examples.

5. After the examples, ensure the pedagogical agent instructs the student to initiate the task. The procedure remains consistent throughout task execution, recording both correct and incorrect responses.

6. If the student needs to listen to a word again, click on the **speaker Icon** for a replay. Note that only one replay per item is allowed.

3. Orange Door: Naming speed and orthographical modules

1. Rapid automatized naming (RAN) task

1. Access the program on the computer. Proceed to the orange door within the program interface. Choose **Naming Speed** to select the task.

2. Permit the pedagogical agent to provide instructional guidance and demonstrate the task. Present an example of the letter RAN subtask of the naming speed module.

3. Ask the students to articulate their answers aloud. If the response in the example is accurate, select the blue button on the screen. If incorrect, click on the red button.

4. Ensure the pedagogical agent provides feedback based on the student's responses.

after the example, ensure that the pedagogical agent initiates the task.

5. Start the task by clicking the left mouse button to begin timing. During the task, a matrix of elements will appear on the screen, depending on the subtask.

6. Ask the student to name aloud and, in order, the elements of the matrix. Simultaneously, document any errors made by the student.

7. For each error, use the right button of the external mouse to register the number of errors. Once the student has finished naming the elements, press the left button again to conclude the timing and complete the task for that subtask. Upon completion of the first subtask (letter RAN), the next subtask (number RAN) automatically appears, following the same procedure. This process continues for the remaining subtasks (color RAN and object RAN).

2. Lexeme and Suffixes task

   1. Access the program on the computer. Navigate to the orange door within the platform interface and click on **Lexemes and Suffixes** to select the task.

   2. Enable the pedagogical agent to offer instructional guidance and task modeling to the student, including an example item where the student uses an external mouse to select an image corresponding to the target word.

   3. Ensure the pedagogical agent provides feedback based on the student's response to the example and then instructs the student to begin the task.

4. Maintain procedural consistency throughout the task execution, prompting the student to select an image corresponding to the target word for each item presented and recording both correct hits and missed responses.

3. Homophone comprehension task

   1. Access the program on the computer. Navigate to the orange door within the platform interface. Click on **Homophone Comprehension** to select the Task.

   2. Enable the pedagogical agent to offer instructional guidance and task modeling to the student. Present an example item, prompting the student to use an external mouse to select an image corresponding to the target word.

   3. Ensure the pedagogical agent provides feedback based on the student's response to the example and instruct the student to begin the task following the example presentation.

   4. Maintain consistency in the procedural steps throughout the task execution. Prompt the student to select an image corresponding to the target word for each item presented.

   5. Record both correct hits and missed responses during task execution.

4. Green Door: Syntactic processing module

   1. Gender task

      1. Open the program on the computer. Navigate to the green door within the program interface. Click on **Gender** to select the task.

2. Allow the pedagogical agent to provide the student with instructions and task modeling. Present two examples for the student to complete, requiring them to click on the corresponding words in each sentence according to their gender agreement using an external mouse.

3. Ensure the pedagogical agent provides feedback based on the student's response to the examples. Following the example presentation, ensure the pedagogical agent guides the student to begin the task.

4. Throughout the task execution, maintain consistency in the procedure. Prompt the student to click on the corresponding words in each sentence according to their gender agreement.

5. Collect both correct and incorrect responses during task execution.

   NOTE: Unlike Spanish, English does not distinguish gender in its grammar.

2. Number task

   1. Open the program on the computer. Navigate to the green door within the program interface. Click on **Number** to select the task.

   2. Allow the pedagogical agent to provide the student with instructions and task modeling. Present two examples for the student to complete, requiring them to click on the corresponding words in each sentence according to their number agreement using an external mouse.

3. Ensure the pedagogical agent provides feedback based on the student's response to the examples. Following the example presentation, ensure the pedagogical agent guides the student to begin the task.

4. Throughout the task execution, maintain consistency in the procedure. Prompt the student to click on the corresponding words in each sentence according to their number agreement.

5. Collect both hits and misses during task execution.

   NOTE: Unlike Spanish, English does not distinguish numbers in its grammar.

3. Functional words task

   1. Open the program on the computer. Navigate to the green door within the program interface. Click on **Functional Words** to select the task.

   2. Allow the pedagogical agent to provide the student with instructions and task modeling. Present two examples for the student to complete, requiring them to click on the corresponding function words in each sentence according to the sentence's context using an external mouse.

   3. Ensure the pedagogical agent provides feedback based on the student's response to the examples. Following the example presentation, ensure the pedagogical agent guides the student to begin the task.

   4. Throughout the task execution, maintain consistency in the procedure. Prompt the

student to click on the corresponding function words in each sentence according to the sentence's context.

5. Collect both correct and incorrect responses during task execution.

4. Grammatical structure task

1. Open the program on the computer. Navigate to the green door within the program interface. Select **Grammatical Structure** to choose the task.

2. Allow the pedagogical agent to provide the student with instructions and task demonstrations. Present two examples for the student to complete, requiring them to click on the appropriate sentence for each picture using an external mouse.

3. Ensure the pedagogical agent provides feedback based on the student's response to the examples. Following the example presentation, ensure the pedagogical agent guides the student to commence the task.

4. Throughout the task execution, maintain consistency in the procedure. Prompt the student to click on the appropriate sentence for each picture using an external mouse.

5. Capture both correct and incorrect responses during task execution.

5. Blue Door: Semantic processing module using reading comprehension task

1. Open the program on the computer. Proceed to the blue door within the program interface. Choose either **The Fruits** for the informative text or **Tino's Getaway** for the narrative text.

2. Allow the pedagogical agent to provide the student with task instructions. Once the type of text has been chosen, display the text on the screen.

3. The student must read the text and memorize the most relevant information. After finishing reading, ask the student to click on the arrow on the screen using the external mouse to indicate completion of the reading and proceed to the next section.

4. Allow the pedagogical agent to instruct the student to read the questions and select the correct answer using the external mouse. The student reads the questions and selects the correct answer accordingly.

3. Data analysis

1. To assess the construct validity and diagnostic accuracy of the multimedia battery in a Spanish-speaking population, use confirmatory factor analysis (CFA) and receiver operating characteristic (ROC) curve analysis.

2. Perform CFA to validate the underlying factor structure of the multimedia battery. This analysis allows to test the hypothesis that the data fit a predefined structure based on theoretical expectations.

3. Evaluate the model fit using several fit indices, including the comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). A good model fit is indicated by CFI and TLI values greater than 0.90, RMSEA values less than 0.08, and SRMR values less than 0.08.

4. To determine the diagnostic accuracy of the multimedia battery, perform ROC curve analysis. This method allows to evaluate the ability of the test to correctly classify individuals with and without reading difficulties. The area under the ROC curve (AUC) provides a measure of the test's overall accuracy. An AUC of 0.5 indicates no diagnostic ability, whereas an AUC of 1.0 indicates perfect diagnostic ability.

5. Identify optimal cutoff points for the battery by analyzing the sensitivity and specificity at various threshold levels.

6. By employing both CFA and ROC curve analysis, perform a comprehensive evaluation of the multimedia battery, confirming its construct validity and diagnostic accuracy in a Spanish-speaking population.

Representative Results

Sample study

The sample included 881 participants from Spain (N = 325), Mexico (N = 169), Guatemala (N = 227), and Chile (N = 160), all of whom were native Spanish speakers. The sample was divided into two groups: 451 in the reading disability (RD) group and 430 in the normally achieving readers (NAR) group. Children with special educational needs-those requiring support and specific educational attention due to sensory impairments, neurological issues, or other conditions-were excluded because these factors are typically used as exclusionary criteria for learning disabilities or severe behavioral disorders, either temporarily or for the duration of their schooling. Participants in the RD group were selected based on criteria such as an IQ equal to or greater than 80, a word reading time measurement percentile above the 75th percentile, a pseudoword reading time measurement percentile above the 75th percentile, or a pseudoword reading accuracy measurement percentile below the 25th percentile.

Similarly, the NAR group was selected based on comparable criteria, with the addition of a PROLEC[36] comprehension task percentile above the 25th percentile. All percentiles were determined according to participants' grade levels. The classification of participants into these groups relied on tasks of word reading and pseudoword reading, each comprising separate blocks. In both blocks, participants were instructed to read aloud the presented verbal stimuli as quickly as possible. The word reading block included 32 stimuli, while the pseudoword reading block contained 48 stimuli. The familiarity of the words was controlled using criteria outlined by Guzman and Jiménez[37]. Measures of hits, errors, and latency times were recorded, with latency times measured from the appearance of each item to the start of the student's vocal response. The word reading block demonstrated high internal consistency (α = .89), while the pseudoword reading block exhibited even greater internal consistency (α = .91). The distribution of students with RD and NAR across different grade levels was as follows: In the second grade, there were 30 NAR and 27 RD students. Transitioning to the third grade, 100 NARs were observed alongside 88 RD students. The fourth grade included 100 NAR and 116 RD students. In the fifth grade, there were 100 NAR and 116 RD students, while in the sixth grade, there were 100 NAR and 104 RD students. To assess age differences in the total sample and within each grade level, one-way ANOVA tests were conducted. The results revealed no significant differences, with all p-values exceeding 0.05. Similarly, the distribution of the gender variable was examined using the chi-square test, with all p values also exceeding 0.05.

Descriptive statistics

Table 1 presents the descriptive statistics (mean and standard deviation) for age and the assessment measures. The data are categorized by gender, distinguishing between

normally achieving readers and those with reading disabilities across various countries. The results showed positive and statistically significant correlations among a large number of assessment measures (see **Table 2**). The table displays correlations among the measures, spanning from weak to strong. Correlations less than 0.3 are considered weak, while those greater than 0.5 are categorized as strong. Specific correlation values are provided alongside their respective strength classifications. For example, weak correlations (r < 0.3) included those of blending-segmentation (r = 0.354) and deletion-homophone comprehension (r = 0.270). Moderate correlations (0.3 ≤ r < 0.5) were observed between the grammatical structure-number (r = 0.463) and the functional words-grammatical structure (r = 0.512). Strong correlations (r ≥ 0.5) are evident for number-gender (r = 0.642) and picture RAN-Color RAN (r = 0.442).

Table 1: Descriptive statistics of the measures. This table provides the mean, standard deviation, and other descriptive statistics for the various measures included in the digital tool. AGE: Age; NAR: Normally Achieving Readers; RD: Reading disability ; SEG: Segmentation; BLN: Blending; ISO: Phoneme isolation; DEL: Phoneme deletion; HOM: Homophone comprehension; R&S: Root & Suffixes; GEN: Gender; NUM: Number; FW: Functional words; GRS: Grammatical Structure; EXT: Expositive Text ; NAT: Narrative Text ; VOC: Voicing of articulation; MOA: Manner of articulation; POA= Place of articulation; DIG: Digit RAN; LET: Letter Ran; COL: Color RAN; PIC: Picture RAN Please click here to download this Table.

Table 2: Correlation coefficients of the measures. This table shows the correlation coefficients between different measures of the digital tool, indicating the strength and direction of the relationships between these measures.

SEG: Segmentation; BLN: Blending; ISO: Phoneme isolation; DEL: Phoneme deletion; HOM: Homophone comprehension; R&S: Root & Suffixes; GEN: Gender; NUM: Number; GRS: Grammatical Structure; FW: Functional Words; EXT: Expositive Text ; NAT: Narrative Text ; VOC: Voicing of articulation; MOA: Manner of articulation; POA= Place of articulation; DIG: Digit RAN; LET: Letter Ran; COL: Color RAN; PIC: Picture RAN. Please click here to download this Table.

Confirmatory factor analysis

CFA was conducted to assess the proposed factor structure of the multimedia battery. The model includes one second-order factor and six latent variables, each representing distinct modules of the multimedia battery. These constructs include phonological awareness (including phonemic segmentation, isolation, blending, and deletion tasks), morpho-orthographic module (involving root and suffixes and homophone comprehension tasks), syntax (comprising gender, number, grammatical structure, and functional word tasks), speech perception (involving voicing, manner of articulation, and place of articulation tasks), reading comprehension (encompassing expositive text and narrative text comprehension tasks), and rapid automated naming (letters, colors, digits, and pictures tasks from the RAN task). To ensure consistency across all tasks in the multimedia battery (ranging from 0 to 1), we used the proportion of maximum scaling (POMS) estimation. POMS scores were computed for the time-based tasks (roots and suffixes)[38].

Statistical analyses and graphical presentations were performed using R version 4.3.1[39], employing the lavaan[40], semTools[41], and ggplot2[42] packages. Model fit was assessed using various goodness-of-fit indices. Although

the chi-square goodness of fit was significant, χ^2(df) = 632.01, p < .001, which suggests a discrepancy between the hypothesized model and the observed data, it is important to note that the chi-square test is sensitive to sample size[43]. Therefore, additional fit indices were considered.

The comparative fit index (CFI) yielded a value of .961, exceeding the commonly accepted threshold of .95 and indicating a good fit to the data. The root mean square error of approximation (RMSEA) was .038, indicating a close fit of the model to the data. The standardized root mean square residual (SRMR) was .034, which was below the recommended value of .05, suggesting an ideal fit. Factor loadings for the items ranged from .36 to .81, indicating significant loadings on their respective factors. The total average variance extracted (AVE) exceeded .50, suggesting adequate convergent validity, and the total composite reliability (CR) was above .80, indicating good internal consistency reliability.

In conclusion, the results of the CFA supported the proposed factor structure, demonstrating good model fit, convergent validity, and reliability. Please refer to the graphical representation of the model in **Figure 1**.
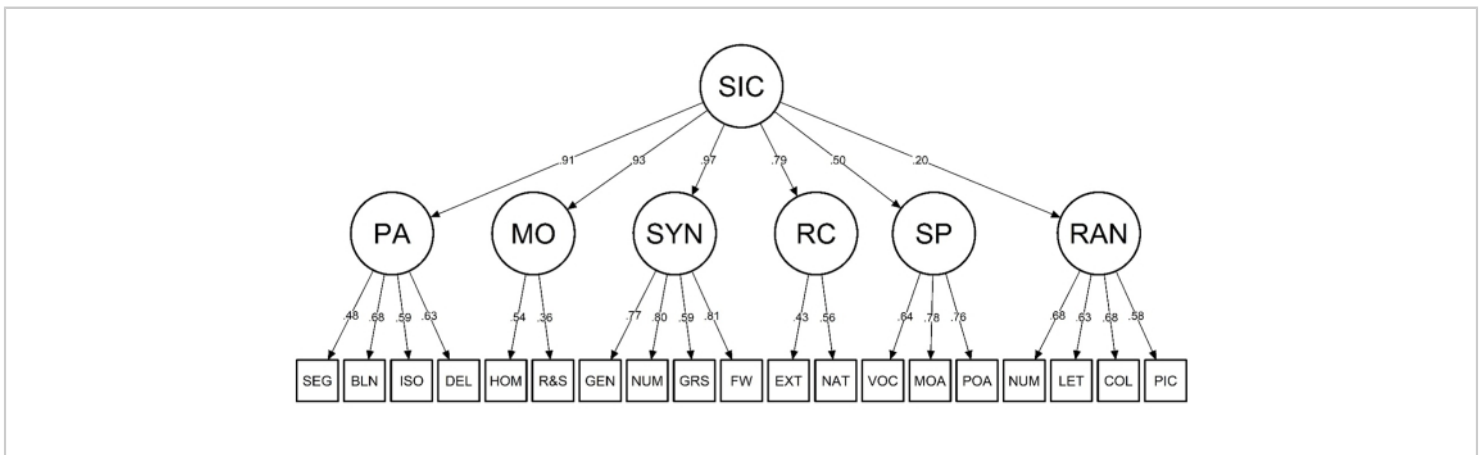

Figure 1: Confirmatory factor analysis. This figure illustrates the results of the confirmatory factor analysis performed on the digital tool, highlighting the factor loadings and the relationships between different cognitive processing measures. Abbreviations: SIC= Sicole-R; PA= Phonological awareness; SEG= Segmentation; BLN= Blending; ISO= Isolation; DEL= Deletion; MO= Morphological processing; HOM= Homophone comprehension; R&S= Root and suffixes; SYN= Syntactic processing; GEN= Gender; NUM= Number; GRS= Grammatical structure; FW= Functional words; RC= Reading comprehension; EXT= Expositive text; NAT= Narrative text; SP= Speech perception; VOC= Voicing of articulation; MOA= Manner of articulation; POA= Place of articulation; RAN= Rapid Automatized Naming; NUM= Number RAN; LET= Letter RAN; COL= Color RAN; PIC= Picture RAN. Please click here to view a larger version of this figure.

Measurement invariance analyses were used to test whether the factor structure of the multimedia battery was stable across genders. Testing for measurement invariance consists of a series of model comparisons that define increasingly stringent equality constraints[44]. We carried out the statistical analyses and followed the same fit model

criteria of the previous CFA. We successively constrained parameters representing the configural, metric (loadings), scalar (intercepts), and strict (residuals) structures[45]. A poor fit in any of these models suggests that the aspect being constrained does not operate consistently for the different groups. The degree of invariance was determined jointly when χ^2_D was >0.05 and ΔCFI was <0.01. A summary of the results of the indices of the models and the differences between them are shown in **Table 3.**

Table 3: Invariance model fit indices. This table presents the fit indices for testing the gender invariance of the model, assessing how well the digital tool maintains consistency in its structure and measurement properties across male and female groups. These indices ensure the tool's reliability and validity by confirming its ability to measure cognitive processes consistently across diverse gender groups. Please click here to download this Table.

The configural model, which constrained only the relative configuration of variables in the model to be the same in both groups, had an adequate fit to the data: $\chi^2(292)$ = 745.970, p<.001, CFI = .964, SRMR = 0.034, RMSEA = .036, 90% CI (.033, .038). The metric invariance model constrained the configuration of variables and all factor loadings to be constant across groups. The fit indices were comparable to those of the configural model: $\chi^2(310)$ = 768.56, p<.001, CFI = .963, SRMR = .037, RMSEA = .036, 90% CI (.033, .038). The invariance of the factor loadings was supported by the nonsignificant difference tests that assessed model similarity: χ^2_D (18) = 26.30, p =.09; ΔCFI = .001. In the scalar invariance model, the configuration, factor loadings, and indicator means/intercepts were constrained to be the same for each group. The fit indices were less than ideal: $\chi^2(322)$ = 787,50, p<.001, CFI = .963, SRMS = .038, RMSEA = .035

(90% CI = .032, .038). The difference tests that evaluated model similarity suggested that there was factorial invariance: $\chi^2_D(12)$=13.00, p = .369; ΔCFI = .001. Finally, in the strict invariance model, the configuration, factor loadings, indicator means/intercepts, and residuals were constrained to be the same for each group. The fit indices were less than ideal: $\chi^2(341)$ =798.20, p<.001, CFI = .961, SRMS = .039, RMSEA = .0035 (90% CI = .032, .038). Strict invariance was supported by the nonsignificant difference tests that assessed model similarity: $\chi^2_D(11)$ = 25.81, p<.001; ΔCFI =.002

Diagnostic accuracy

To evaluate the accuracy and discriminative capacity of the multimedia battery, receiver operating characteristic (ROC) analyses were conducted. ROC analysis aids in determining the optimal threshold (cutoff value) for a continuous-scale assessment test, balancing sensitivity (ability to correctly identify true positives) and specificity (ability to correctly identify true negatives). Additionally, ROC analysis was used to assess the ability of the test to discriminate between the RD and NAR groups. To carry out the analysis, Z scores were calculated per grade for each task of the multimedia battery. The omnibus score consisted of the sum of the zeta scores. Statistical analyses and graphical presentations were carried out using R version 4.3.1[41]. The pROC[46] and ggplot2[42] packages were used. In terms of diagnostic accuracy, the multimedia battery exhibited an area under the curve (AUC) of 9439.8 [95% CI: 93.31%-96.24% (DeLong)] and a sensitivity of 91.0 (**Figure 2**). The ROC curves by grade showed the following indices: 2° grade, AUC= 96.195% [CI: 91.54%-100% (DeLong)], Se= .96, Sp=.90; 3° grade, AUC= 95.3, [95% CI: 92.43%-98.18% (DeLong)], Se=75.70, Sp= 72.34; 4° grade, AUC=93.4 [95% CI: 90.2%-96.66% (DeLong)], Se=.92, Sp=.84; 5° grade, AUC=95.9 [95% CI: 93.11%-98.75% (DeLong)], Se=.90, Sp=.95; 6° grade,

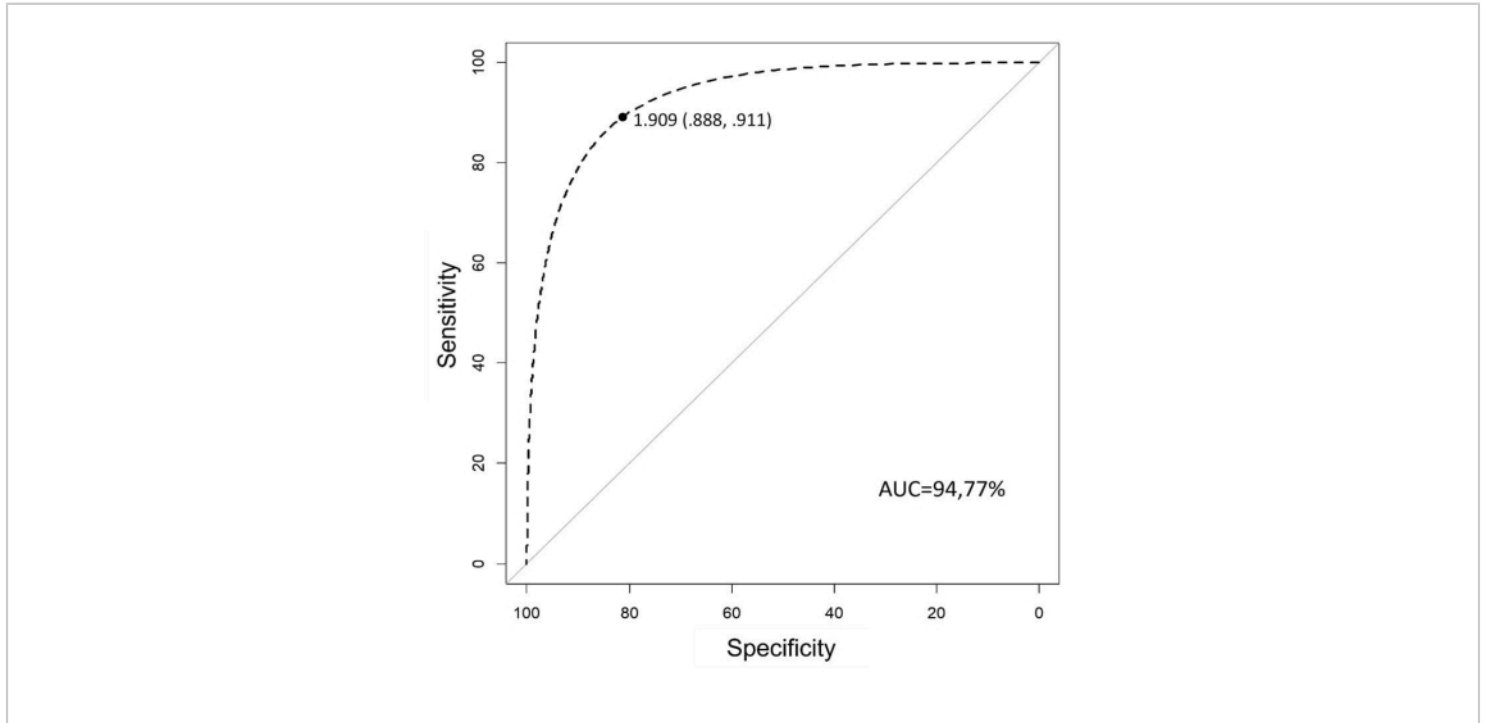AUC=94.4 [95% CI: 91.11%-97.69% (DeLong)], Se=.92, Sp=.91.

Figure 2: Curve ROC analysis. This figure presents the ROC curve analysis, which shows the diagnostic accuracy of the digital tool by plotting the true positive rate against the false positive rate at various threshold settings. The units for the x-axis of the ROC curve are specificity, and the units for the y-axis are sensitivity. Abbreviations: ROC= receiver operating characteristic curve; AUC= area under the curve. Please click here to view a larger version of this figure.

Discussion

In this study, confirmatory factor analysis (CFA) was employed to evaluate the factor structure of the Sicole-R battery, comprising one second-order factor and six latent variables representing different modules. The results indicated good model fit, convergent validity, and reliability, confirming the efficacy of the battery in assessing a comprehensive set of cognitive and reading skills that are critical for individuals with dyslexia. Importantly, the consistent performance of the digital tool across diverse demographic groups within Spain, Mexico, Guatemala, and Chile suggests

its potential utility in educational and clinical settings across various Spanish-speaking regions.

This study contributes to literature by demonstrating how multimedia batteries can be used to effectively identify specific cognitive deficits associated with dyslexia. The battery's ability to pinpoint specific areas of difficulty, such as phonological awareness, speech perception, naming speed, orthographical processing, syntactic processing, and reading comprehension, enables tailored interventions to address individual student needs effectively. Moreover, the demonstrated consistency of the battery across genders

underscores its utility in promoting equitable assessment practices, thereby supporting inclusive education initiatives aimed at providing equal opportunities for all students.

Although previous tools with similar characteristics were designed for use in the English language, such as the Dyslexia Early Screening Test (DEST) proposed by Fawcett and Nicolson[17], which assesses multiple domains, this study demonstrated the applicability and effectiveness of the multimedia battery in a different linguistic context. Throughout the analysis, several critical steps were undertaken within the protocol to ensure the accuracy and validity of the results. These steps included data preprocessing, model specification, and parameter estimation, aligning with studies such as Hautala et al.[19] that have demonstrated the reliability of computer-based tools in identifying students with reading difficulties. Additionally, other studies, such as those by Protopapas et al.[20] and Pennington et al.[12], have explored the implementation of computer-based screening tools for dyslexia identification, underscoring the importance of harnessing technology to enhance assessment strategies. Furthermore, the current study aligns with recent advancements in dyslexia research and psychometric assessment, emphasizing the relevance of the battery in contemporary contexts. Studies conducted by Jiménez et al.[23] and Guzmán et al.[24] have highlighted the critical need for robust assessment tools tailored to diverse linguistic backgrounds, underscoring Sicole-R's efficacy in filling this gap within Spanish-speaking populations. Moreover, the integration of digital technologies in dyslexia assessment, as discussed by Rauschenberg et al.[6], reinforces the methodological rigor demonstrated in this study's approach to factor analysis and validation.

Despite the methodological rigor employed in this study, it is important to acknowledge certain limitations inherent in the analysis. One limitation concerns the generalizability of the findings, as the sample consisted of a specific demographic group, potentially limiting the applicability of the results to broader populations. Additionally, while efforts were made to ensure measurement invariance across genders, other demographic factors, such as age or socioeconomic status, were not explicitly accounted for in the analysis, which may have influenced the results.

Future research directions could explore its adaptation to different Spanish dialects and its integration into broader educational frameworks to enhance identification strategies. Moving forward, future research should aim to address the limitations identified in this study, such as exploring the utility of the multimedia battery in diverse populations and investigating its longitudinal validity and predictive validity in predicting long-term educational outcomes. Additionally, a promising future line of research has been highlighted in a recent review by Jin et al.[47], suggesting the incorporation of neurobiological technology. They conducted a comprehensive review of scientific databases, selecting studies published between 2018 and 2023. Their findings suggest that neurobiological technology assessment is emerging as a promising trend in advancing the diagnosis of dyslexia.

In conclusion, the findings of this study highlight the multimedia battery as a valuable tool for educators, ultimately facilitating more effective support and improved outcomes for students with dyslexia. By incorporating these contemporary insights, the findings contribute to the evolving discourse on dyslexia assessment, offering practitioners and researchers updated perspectives on effective diagnostic

strategy frameworks. To our knowledge, this study is among the first to validate a comprehensive multimedia battery tailored specifically for Spanish-speaking populations, filling a critical gap in the current diagnostic tools available for dyslexia.

Disclosures

The authors listed above certify that there are no financial interests or other conflicts of interest regarding the present study.

References

1. World Federation of Neurology. *Report of research group on Dyslexia and world illiteracy.* Dallas (1968).

2. Diamanti, V., Goulandris, N., Campbell, R., Protopapas, A. Dyslexia profiles across orthographies differing in transparency: An evaluation of theoretical predictions contrasting English and Greek. *Sci Stud Reading.* **22** (1), 55-69 (2018).

3. Shaywitz, S. *Overcoming dyslexia: A new and complete science-based program for reading problems at any level.* Alfred A. Knopf (2003).

4. Heikki Lyytinen, H., Ronimus, M., Alanko, A., Poikkeus, N. M., Taanila, M. Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. *Nordic Psychol.* **59** (2), 109-126 (2007).

5. Tobia, V., Marzocchi, G. M. Cognitive profiles of Italian children with developmental dyslexia. *Reading Res Quart.* **49***(4), 437.*-452 (2014).

6. Rauschenberger, M., Baeza-Yates, R., Rello, L. A universal screening tool for dyslexia by a web-game and machine learning. *Front Comput Sci.* **3**, 628634 (2022).

7. Ekhsan, H.M., Ahmad, S.Z., Halim, S.A., Hamid, J.N., Mansor, N.H. The implementation of interactive multimedia in early screening of dyslexia. *Int Conf Innov Mgmt Tech Res.* 566-569 (2012).

8. Jiménez, J.E. *Dyslexia in Spanish. Prevalence and cognitive, cultural, family and biological indicators.* Ediciones Pirámide, Madrid (2012).

9. Hatcher, J., Snowling, M. J. The phonological representations hypothesis of dyslexia: From theory to practice. *Dyslexia and Literacy, Theory and Practice.* Chichester, John Wiley & Sons (2002).

10. Goswami, U. A temporal sampling framework for developmental dyslexia. *Trends Cogn Sci.* **15** (1), 3-10 (2011).

11. Stein, J. F. The magnocellular theory of developmental dyslexia. *Dyslexia.* **7** (1), 12-36 (2001).

12. Nicolson, R. I., Fawcett, A. J. Development of Dyslexia: The delayed neural commitment framework. *Front Behav Neurosci.* **21** (13), 112 (2001).

13. Pennington, B. F. et al. Individual prediction of dyslexia by single versus multiple deficit models. *J Abnorm Psychol.* **121** (1), 212-224 (2012).

14. Ring, J., Black, J. L. The multiple deficit model of dyslexia: what does it mean for identification and intervention? *Ann Dyslexia.* **68** (2), 104-125 (2018).

15. Soriano-Ferrer, M., Nievas-Cazorla, F., Sánchez-López, P., Félix-Mateo, V., González-Torre, J. A. Reading-related cognitive deficits in Spanish developmental Dyslexia. *Procedia - Social Behav Sci.* **132**, 3-9 (2014).

16. Zygouris, N.C. et al. *The implementation of a web application for screening children with Dyslexia.* International Conference on Interactive Collaborative Learning (2016).

17. Snowling, M. J., Hulme, C. *The Science of Reading: A Handbook.* Wiley-Blackwell (2021).

18. Politi-Georgousi, S., Drigas, A. Mobile applications, an emerging powerful tool for Dyslexia screening and intervention: A systematic literature review. *Int J Interact Mobile Technol.* **14** (18), 4-17 (2020).

19. Fawcett, A. J., Nicolson, R. I. The Dyslexia early screening test. *Irish J Psychol.* **16** (3), 248-259 (1995).

20. Zygouris, N. C. et al. *The implementation of a web application for screening children with dyslexia. Interactive Collaborative Learning.* Advances in Intelligent Systems and Computing, Springer, Cham (2017).

21. Hautala, J. et al. Identification of reading difficulties by a digital game-based assessment technology. *J Edu Comput Res.* **58** (5), 1003-1028 (2020).

22. Ahmad, N., Rehman, M.B., El Hassan, H.M., Ahmad, I., Rashid M. An efficient machine learning-based feature optimization model for the detection of dyslexia. *Comput Intell Neurosci.* **9**, 8491753 (2022).

23. Rello, L., Baeza-Yates, R., Ali, A., Bigham, J.P., Serra M. Predicting risk of dyslexia with an online gamified test. *PLoS One.* **15** (12), e0241687 (2020).

24. Dębska, A. et al. The cognitive basis of dyslexia in school-aged children: A multiple case study in a transparent orthography. *Dev Sci.* **25** (2), e13173 (2022).

25. Jiménez, J. E. et al. Discriminant validity of the Sicole-R-Primaria Multimedia Battery for the evaluation of cognitive processes associated with dyslexia. *Revista de Investigación Educativa.* **27** (1), 49-71 (2009).

26. Guzmán, R. et al. Assessment of naming speed in reading learning difficulties. *Psicothema.* **16** (3), 442-447 (2004).

27. Jiménez, J. E. et al. Is the deficit in phonological awareness better explained in terms of task differences or effects of syllable structure? *Appl Psycholinguist.* **26***(2), 267.*-283 (2005).

28. Ortiz, R. et al. Locus and nature of perceptual phonological deficit in Spanish children with reading disabilities. *J Learn Disabil.* **40** (1), 80-92 (2007).

29. Ortiz, R. et al. Development of speech perception in children with dyslexia. *Psicothema.* **20** (4), 678-683 (2008).

30. Jiménez, J. E. et al. The Double-deficit hypothesis in Spanish developmental dyslexia. *Topics Lang Disorders.* **28** (1), 46-60 (2008).

31. Jiménez, J.E., García de la Cadena, C. Learning disabilities in Guatemala and Spain: A cross-national study of the prevalence and cognitive processes associated with reading and spelling disabilities. *Learning Disabilities Res Pract.* **22,** 161-169 (2007).

32. Jiménez, J. E., de la Cadena, C. G., Siegel, L. S., O'Shanahan, I., García, E. Gender ratio and cognitive profiles in dyslexia: a cross-national study. *Read Writ.* **24**, 729-747 (2011).

33. Rodrigo, M. et al. Assessment of orthographic processing in Spanish children with dyslexia: the role of lexical and sublexical units.*Revista Electrónica de Investigación Psicoeducativa y Psicopedagógica.* **2** (2), 105-126 (2004).

34. Jiménez, J. E. et al. Evaluation of syntactic-semantic processing in developmental dyslexia. *Elect J Res Edu Psychol.* **2** (2), 127-142 (2004).

35. Jiménez, J.E., Rodríguez, C., Ramírez, G. Spanish developmental dyslexia: prevalence, cognitive profile, and home literacy experiences. *J Exp Child Psychol.* **103** (2), 167-85 (2009).

36. Cuetos, F., Rodríguez, B., Ruano, R., Arribas, A.D. *PROLEC-R. Primary Reading Processes Assessment Battery - Revised.* TEA Ediciones (2005).

37. Guzmán, R., Jiménez, J. E. Normative study on psycholinguistic parameters in children aged 6 to 8 years: Subjective familiarity. *Cognitiva.* **2**, 153-191 (2001).

38. Cohen, P., Cohen, J., Aiken, L., West, S.G. The problem of units and the circumstance for POMP. *Multivariate Behav Res.* **34**, 315-346 (1999).

39. R Core Team. *_R: A Language and Environment for Statistical Computing_.* R Foundation for Statistical Computing, Vienna, Austria (2023).

40. Rosseel, Y. lavaan: An R package for structural equation modeling. *J Stat Software.* **48** (2), 1-36 (2012).

41. Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y. *semTools: Useful tools for structural equation modeling.* R package version 0.5-6 (2022).

42. Wickham, H. *ggplot2: Elegant graphics for data analysis.* Springer-Verlag, New York (2016).

43. Byrne, B. M. *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* Third Edition. Taylor & Francis Group (2016).

44. Vandenberg, R. J., Lance, C. E. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Org Res Methods.* **3** (1), 4-70 (2000).

45. Hirschfeld, G., von Brachel, Ruth. Multiple-Group confirmatory factor analysis in R - A tutorial in measurement invariance with continuous and ordinal indicators. *Pract Assess Res Eval.* **19** (7), 1-12 (2014).

46. Sachs, M. C. plotROC: A Tool for Plotting ROC Curves. *J Stat Softw. 79, .*2 (2017).

47. Jing, W., Yu, E. L. X., Motevalli, S. A comparison between cognitive assessment and neurobiology technology assessment of dyslexia: A literature review. *Int J Acad Res Busi Soc Sci.* **13** (4), 976 - 986 (2023).