Spanish Journal of Psychology (2015), **18**, e18, 1–12.

doi:10.1017/sjp.2015.15

CAMBRIDGE UNIVERSITY PRESS

The Loleva Oral and Written Language Test: Psychometric Properties

Manuel Peralbo[1], María Ángeles Mayor[2], Begoña Zubiaz[2], Alicia Risso[1], María Luz Fernández-Amado[3] and Alejandro Tuñas[4]

[1] *Universidad de A Coruña (Spain)*
[2] *Universidad de Salamanca (Spain)*
[3] *Colegio San Juan Bosco (Spain)*
[4] *Altia Consultores (Spain)*

Abstract. *LolEva,* a computerized test for ages 3 to 8 years old, identifies issues in the development of skills that can lead to reading acquisition difficulties. Its structure captures two distinct areas: Phonological Awareness (PA, seven subtests: rhyme, identification-addition-omission of syllable and phoneme at the beginning and end of a word), and Initial Reading Competence (IRC, six subtests: reading uppercase and lowercase letters, simple words, complex words, and pseudowords, and word segmentation). With results collected in a sample of 341 children with the target ages and attending public or private schools, the alpha coefficient was .94 for PA, and .92 for IRC. Factor analysis indicated three factors are present (performance on PA and IRC, and word reading times), together explaining 75% of variance, providing evidence to support the construct validity of the test. On the other hand, analysis of variance showed significant differences for year-in-school variable for PA subscale, $F(4, 336) = 191.385$, $p < .001$, $\eta^2_p = .695$, $1–\beta = 1.0$, as well as for IRC subscale, both in number of correct answers, which increased as schooling progressed: $F(4, 336) = 197.897$, $p < .001$, $\eta^2_p = .702$, $1–\beta = 1.0$, and task completion time, which decreased as education progressed: $F(4, 335) = 47.048$, $p < .001$, $\eta^2_p = .360$, $1–\beta = 1.0$. Also, PA repeated measures analysis revealed that was easier Identification than Addition and Omission , $F(2, 672) = 31.639$, $p < .001$, $\eta^2_p = .086$, $1–\beta = 1.0$, syllable-related tasks than phoneme-related task, $F(1, 336) = 229.000$, $p < .001$, $\eta^2_p = .405$, $1–\beta = 1.0$, and syllable or phoneme at the end of the word than at the beginning, $F(1, 336) = 59.201$, $p < .001$, $\eta^2_p = .150$, $1–\beta = 1.0$. Moreover, all items were examined and indexes of difficulty and discrimination were obtained.

Received 27 January 2014; Revised 31 July 2014; Accepted 13 August 2014

Keywords: phonological awareness, reading skills, computer assisted testing, Psychometrics.

This paper presents a psychometric analysis of the *Lenguaje Oral, Lenguaje Escrito: Evaluación* (*LolEva*) [Oral & Written Language: Assessment] (Mayor, Fernández, Tuñas, Zubiauz, & Durán, 2012; Peralbo, Brenlla, García, Barca, & Mayor, 2012). Taking a psycholinguistic perspective, the test is built on the assumption that an interacting set of regular, continuous processes brings about language comprehension and use, both oral and written. Metalinguistic processes refer to a person's awareness of the component parts of words, which seems to facilitate initial reading acquisition in languages with transparent as well as non-transparent orthographies (Alegría, 2006; Caravolas, Lervåg, Defior, Málková, & Hulme, 2013). Phonological awareness has been defined as consciously knowing that words have several constituent units of sound (Gillam & van Kleecl, 1996), and being able to reflect on and manipulate the subunits of speech: syllables, intrasyllabic units, and phonemes (Anthony et al., 2011; Morais, 1991; Treiman, 1991). Morais (1991) put it thusly:

> Phonological awareness is a special kind of phonological knowledge. It differs from the phonological knowledge used in comprehending and producing language by the fact that it refers to conscious representations of the phonological properties and constituents of speech. Indeed, this definition is a loose one unless we specify the criteria by which a phonological representation can be said to be conscious. (p. 34)

Not all knowledge can be easily verbalized, so using recognition or free response tasks, for example, is justified. Furthermore, researchers are not in complete agreement about how phonological awareness develops in different languages, and whether the level of PA a person attains in one language can predict his or her PA in a second language.

This study, however, is based on several assumptions with ample empirical support (Alegría, 2006;

Correspondence concerning this article should be addressed to Alicia Risso. Departamento de Psicología. Campus de Elviña, s/n. Universidad de A Coruña. 15071. A Coruña (Spain). Phone: +34–981167000.

E-mail: alicia.risso@udc.es

Brady & Shankweiler, 1991; Defior, 1996; Jiménez & Ortiz, 1995; Stanovich, 2000):

a. In the development of phonological skills, there is a gradual progression beginning with the ability to manipulate words, followed by syllable manipulation, and finally phoneme manipulation.
b. The position within the word of the unit being manipulated has an effect according to which it is harder to manipulate units at the end of the word that at the beginning, at least on certain types of tasks.
c. The type of syllable structure (CV, VC, CCV, etc.) may influence task complexity.
d. Phonological skills can improve with training, even at a very young age.
e. Phonological skills predict successful initial reading acquisition.
f. There is a reciprocal relationship between phonological awareness and reading.

It has been demonstrated that phonological awareness plays a critical role in learning to read (Bryant & Bradley, 1985; Caravolas et al., 2013; Carroll, Snowling, Stevenson, & Hulme, 2003; Márquez & de la Osa, 2003), and PA assessment has become increasingly necessary for professionals who guide and implement clinical and educational interventions. Utilizing standardized tests adapted for Spanish populations, PA assessment has mostly used a paper-and-pencil format. Examples include the *Prueba de Segmentación Lingüística* (PSL) [Word Segmentation Test] (Jiménez & Ortiz; 1995), *Batería de Evaluación de los Procesos Lectores, Revisada* (PROLEC-R) [Battery of Reading Processes Tests, Revised] (Cuetos, Rodríguez, Ruano, & Arribas, 2007), and the *Test de Análisis de la LectoEscritura* (TALE-2000) [Test of Reading-Writing Analysis] (Toro, Cervera, & Urío, 2000). Nevertheless, ICT (Information and Communication Technology) is only just beginning to be used in developing computerized instruments, instruments like the LolEva (the subject of the present study), *SICOLE-R-Primaria* [SICOLE-R-Primary] (Jiménez et al., 2007), or the *Prueba Informatizada de Habilidades Metafonológicas* [Computerized Test of Metaphonological Skills] (Carrillo & Marín, 1996). To create this sort of computerized test would help complete the current arsenal of tests for the earliest levels of childhood education, covering a wide time span so that phonological awareness difficulties can be identified at an early age and timely intervention can be applied to facilitate the initial learning-to-read process.

A requirement of any assessment tool is that its psychometric properties guarantee it for use and lend credibility to its results (Muñiz, 1999). In terms of validity, the abundant literature to date has confirmed the LolEva's content validity, and its internal structure has emerged from research results. Jiménez and Ortiz (1995) posit that a person's level of phonological awareness can be established according to task difficulty, which varies depending on the linguistic, analytical, and memory aptitude it demands. Similarly, some authors (Leong, 1991; Morais, 1991) draw a distinction between classification and pairing tasks, and segmentation tasks (which require the production or manipulation of isolated elements), arguing that classification tasks are easier. Adams (1990) has proposed as many as five difficulty levels on tasks that measure phonological awareness. In order of difficulty from lowest to highest, the tasks are the following: i. Remember familiar rhymes; ii. Recognize and classify rhyme and alliteration patterns in words; iii. Blend syllables into words, or split one syllable component from the rest (for example, isolate the first phoneme); iv. Word segmentation into phonemes; v. Add, omit, or reverse phonemes, and identify the resulting word or pseudoword.

In their influential study, Anthony et al. (2011) attributed developmental differences between 3 and 6 years of age to the relative influence of task complexity. In their study, items that included the same cognitive operations (addition and omission), or had the same response format (free response or multiple choice) shared a certain amount of variance regardless of the structural segment of the affected word (word, syllable, or phoneme). According to those authors' data, word structure in Spanish does not seem to impact item difficulty. They posit the following developmental sequence of phonological awareness skills (which they believe are one-dimensional). First, children become able to blend two words to form a new one. Next, they learn to connect syllables and form words on their own. Ultimately, the most developed among them can omit sounds from words to create new words. In their study, multiple-choice addition tasks were easier than free-response addition tasks, which were easier than free-response omission tasks. The same sequence was observed regardless of the structural level of the word analyzed. In other words, good results could be obtained if one assesses level of phonological development by selecting just one word structure (syllable or phoneme), and manipulating only task complexity. In Spanish however, as described above, task difficulty seems to stem from other factors, which the aforementioned study ruled out as discriminant (e.g. the type of segment affected, and its position within the word).

Tasks to evaluate reading processes are based on the psycholinguistic processes they involve: from perceptual, linguistic letter recognition; to dual-route lexical access (direct or phonological); and syntactic and semantic analysis. This series of constituent processes is present in practically every test that measures reading processes, and is a point of reference for the LolEva scale of initial reading competence.

With that in mind, this study's objective is to determine and evaluate the main psychometric properties of the LolEva and its constituent parts, specifically: to determine the reliability of total and partial scores (by subscale and subsample), conduct item analysis, and identify the underlying factor structure and, comparing it to the one used when the instrument was created, take a measure of construct validity.

Method

Participants

Participants in this study were 341 students, from the first year in Early Childhood Education (ECE) [*Educación Infantil*] to the second year in Primary Education (PE) [*Educación Primaria*], from public and state-subsidized schools in A Coruña and Salamanca, Spain. They were all 3 to 8 years old, and both sexes were represented (49.6% girls). No one was repeating a grade. The only exclusion criterion was exhibiting severe developmental alteration that could affect comprehension or task performance, but no such case was found. The children were authorized to participate by their parents through an informed consent procedure.

Materials

The computerized LolEva test was employed, designed to identify difficulties in phonological skills development that can lead to problems learning to read. It can be administered to an age range of 3 to 8 years old, and it consists of two subscales: Phonological Awareness (PA) and Initial Reading Competence (IRC). The first is made up of seven tasks with 10 items each: rhyme identification and identification-addition-omission of syllable or phoneme. On half the latter items, the objective syllable or phoneme is positioned at the beginning of the word, and on the other half, it is at the end. The IRC subscale covers six tasks: reading uppercase and lowercase letters (29 items each), reading *simple* words (simple grapheme-phoneme agreement), reading *complex* words (complex grapheme-phoneme agreement; e.g. the Spanish word *quien*), reading pseudowords (10 items each), and splitting sentences into words (*laovejadalana* [thesheepgiveswool]) (5 items). On both subscales, the number of correct responses is recorded, and on the IRC, the time elapsed before responding correctly is measured as well. All instructions and examples are provided in audiovisual format, as determined in advance[1].

[1]Examples of the various tasks can be found at www.loleweb.com--> Loleva--> Vídeo de demostración.

Procedure

The test was administered on an individual basis in classrooms where participants could complete it with no outside interference. The items were presented on a laptop computer (controlled by the test proctor), starting with two examples. Each respondent had to solve at least one sample question on his or her own before the test could be administered. The test took between 40 and 50 minutes to complete. Data collection was carried out during the second semester of the school year in March, April, and May.The data were analyzed using IBM SPSS Statistics 21.

Results

Reliability

To ensure the quality of this measure, we first analyzed the reliability of scores on its subscales, obtaining an alpha coefficient of .94 for PA with all 7 tasks taken into account (Rhyme Recognition, Syllable Identification, Phoneme Identification, Syllable Addition, Phoneme Addition, Syllable Omission, Phoneme Omission), and an alpha coefficient of .96 when the parts corresponding to position (beginning vs. end of word) were considered separately. Scores on the IRC subscale, which included six tasks (Reading Uppercase Letters, Lowercase Letters, Simple words, Complex words, Pseudowords, and Word Separation), had an alpha coefficient of .92. Eliminating any particular item from the set did not produce meaningful differences in the coefficients listed above.

PA Subscale Analysis

The following data inspection aimed to analyze the relationship between results on this subscale and the Year in School variable (Table 1), because in the developmental range covered by the test, having completed different years in school implies having developed different skills that have bearing on what the test measures.

First, a one-way analysis of variance was conducted on total scale scores, with results indicating significant differences and effect size showing that Year in School predicted 70% of variability in PA: $F(4, 336) = 191.385$, $p < .001$, $\eta^2_p = .695$, $1–\beta = 1.0$. However, given the unequal size of year-in-school subsamples, and the inequality of variance, Levène (4.336) = 11.614, $p < .001$, and bearing in mind that F is especially vulnerable when those two conditions apply (Howell, 2013), a robust test of equality of means was conducted. Its results, too, suggest significant differences: Welch (4, 157.662) = 554.676, $p < .001$.

As Table 1 conveys, participants' average number of correct responses increased with year in school.

https://doi.org/10.1017/sjp.2015.15 Published online by Cambridge University Press

Table 1. *Average Proportion of Correct Responses on Each of the Test's Phonological Awareness Tasks in Each School Year*

		1st ECE	2nd ECE	3rd ECE	1st PE	2nd PE
		(n = 39)	(n = 50)	(n = 91)	(n = 74)	(n = 87)
Task		*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Identification	R	.37 (.18)	.57 (.18)	.47 (.31)	.72 (.25)	.80 (.22)
	FS	.33 (.31)	.63 (.25)	.76 (.24)	.89 (.17)	.84 (.22)
	LS	.12 (.19)	.47 (.24)	.70 (.31)	.81 (.25)	.80 (.25)
	FP	.00 (.00)	.04 (.20)	.61 (.30)	.77 (.29)	.80 (.26)
	LP	.00 (.00)	.02 (.14)	.58 (.32)	.70 (.32)	.75 (.30)
Addition	FS	.00 (.00)	.13 (.26)	.51 (.34)	.74 (.29)	.77 (.29)
	LS	.01 (.03)	.26 (.42)	.76 (.30)	.87 (.27)	.91 (.19)
	FP	.00 (.00)	.00 (.00)	.49 (.36)	.60 (.35)	.68 (.31)
	LP	.00 (.00)	.00 (.00)	.69 (.34)	.86 (.26)	.90 (.23)
Omission	FS	.00 (.00)	.18 (.37)	.62 (.38)	.83 (.28)	.89 (.25)
	LS	.01 (.06)	.23 (.39)	.67 (.36)	.86 (.24)	.85 (.22)
	FP	.00 (.00)	.00 (.00)	.37 (.35)	.68 (.31)	.71 (.32)
	LP	.00 (.00)	.00 (.00)	.62 (.37)	.88 (.27)	.88 (.27)
Total		.09 (.04)	.22 (.13)	.59 (.21)	.78 (.19)	.81 (.19)

Note: R = rhyme; FS = first syllable; LS = last syllable; FP = first phoneme; LP = last phoneme.

To ascertain which specific differences were important, we used the Games-Howell test for multiple comparisons, which is considered to have greater power (smaller confidence intervals) than other tests in cases of unequal samples or variances, and can also control FWER (Cardinal & Aiken, 2006; Kirk, 2013; Zimmerman, 2004). Our post-hoc comparison revealed that differences did not occur between all years in school; rather, four subsets were identified: one for each level of ECE, and a single homogenous subgroup combining the two levels of PE. In other words, all differences in means turned out to be statistically significant where $p < .001$, except for the first and second years of PE: difference = –.32, p = .82.

Item Difficulty Indexes (DIs = p-values), too, showed this pattern of similarity and difference between students' results in different years in school. DIs were calculated according to how they are defined within the psychometrics literature: the proportion of people in the sample who answered correctly. It has been said on numerous occasions that it would be more appropriate to refer to this as an item "easiness" index, because the larger it is, the easier the item. With that in mind, this paper will present DIs using quantitative measures as well as the usual qualitative measures from Classical Test Theory.

As Figure 1 shows, for ECE first-years, all the tasks were either difficult (DI in the second quartile, not counting the interval around Q_{50}: .44 ≥ DI ≥ .25) or very difficult (DI in the first quartile: DI ≤ .25), to the point that they only managed to solve the Rhyme Recognition and Syllable Identification tasks. For ECE second-years, however, those last two tasks were easy (DI in the third quartile not counting the interval around Q_{50}: DI ≥ .55) and the rest were very difficult (DI in the fourth quartile: DI ≥ .75). However, the proportion of students able to solve the tasks rose considerably. For ECE third-years, in contrast, most items were easy, but Rhyme Recognition and Phoneme Omission had medium task difficulty (.54 DI ≥ .45). In contrast to the above, first and second-year PE students scarcely differed from one another, although all subtests were very easy for second-years (DI ≥. 75), and that was not the case for first-years.

Regarding item difficulty indexes, we observed that task difficulty generally decreased as year in school progressed, and for all years in school, phoneme-involved tasks were more difficult than syllable-involved tasks.

On another note, the test structure enabled us to construct a within-subjects model to analyze task type (identification/addition/omission), objective type (syllable/phoneme), and objective position (beginning/end), and how they relate to the year-in-school variable. Repeated measures analysis revealed significant main effects, significant interaction effects (Table 2), and a significant between-subjects effect of the year-in-school variable (academic level): $F(1, 4) = 2073.876$, $p < .001$, $\eta^2_p = .708$, $1–\beta = 1.0$.

Furthermore, making multiple comparisons and graphically analyzing student profiles, we observed the following main effects (Figure 2): (1) regarding type of task, identification was easier than Addition

https://doi.org/10.1017/sjp.2015.15 Published online by Cambridge University Press

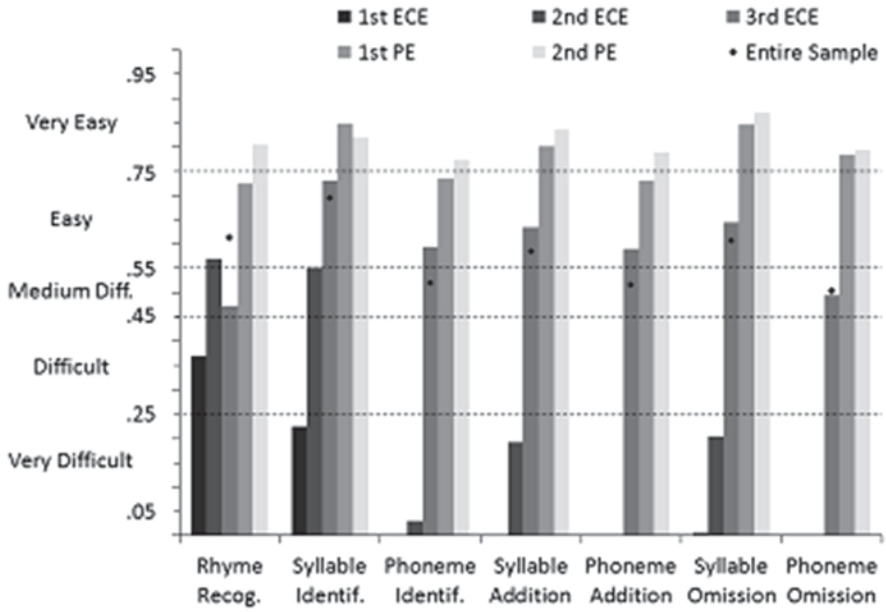

Figure 1. The Phonological Awareness Tasks' Difficulty Indexes.

and Omission, but their difficulty tended to equalize as year in school progressed; (2) about objetive, syllable-related tasks were easier than phoneme-related tasks at all academic levels; (3) regarding objective position, once the ability to solve these tasks had been generally acquired (starting in the third year of ECE), they were easier if the syllable or phoneme was at the end of the word than if it was at the beginning.

Table 2. *Tests of Within-subjects Effects on the Test's Phonological Awareness Tasks*

Source	*F*	*df*	η^2_p	1-β [a]
Task Type [b]	31.639***	2	.086	1.000
Task Type x Year in School [b]	11.129***	8	.117	1.000
Error (Task Type) [b]		672		
Syllable-Phoneme [b]	229.000***	1	.405	1.000
Syllable-Phoneme x Year in School [b]	26.188***	4	.238	1.000
Error (Syllable-Phoneme) [b]		336		
First-Last [b]	59.201***	1	.150	1.000
First-Last x Year in School [b]	18.851***	4	.183	1.000
Error (First-Last) [b]		336		
Task Type x Syllable-Phoneme [c]	35.806***	1.848	.096	1.000
Task Type x Syllable-Phoneme x Year in School [c]	9.664***	7.394	.103	1.000
Error (Task Type x Syllable-Phoneme) [c]		621.060		
Task Type x First-Last [c]	91.001***	1.879	.213	1.000
Task Type x First-Last x Year in School [c]	2.374*	7.516	.027	.879
Error (Task Type x First-Last) [c]		631.367		
Syllable-Phoneme x First-Last [b]	22.994***	1	.064	.998
Syllable-Phoneme x First-Last x Year in School [b]	2.526*	4	.029	.714
Error (Syllable-Phoneme x First-Last) [b]		336		
Task Type x Syllable-Phoneme x First-Last [c]	7.153***	1.986	.021	.931
Task Type x Syllable-Phoneme x First-Last x Year in School [c]	7.892***	7.943	.086	1.000
Error (Task Type x Syllable-Phoneme x First-Last) [c]		667.221		

[a]Calculated for alpha = .05.
[b]Significance of Mauchly's *W* > .05 → Sphericity assumption.
[c]Significance of Mauchly's $W \leq .05$, Epsilon > .70, and $n > k + 10$ → Huynh-Feldt correction for degrees of freedom (Collier, Baker, Mandeville, & Hayes, 1967; Field, 2005; Maxwell & Delaney, 2004).
*p <.001. **p < .05.

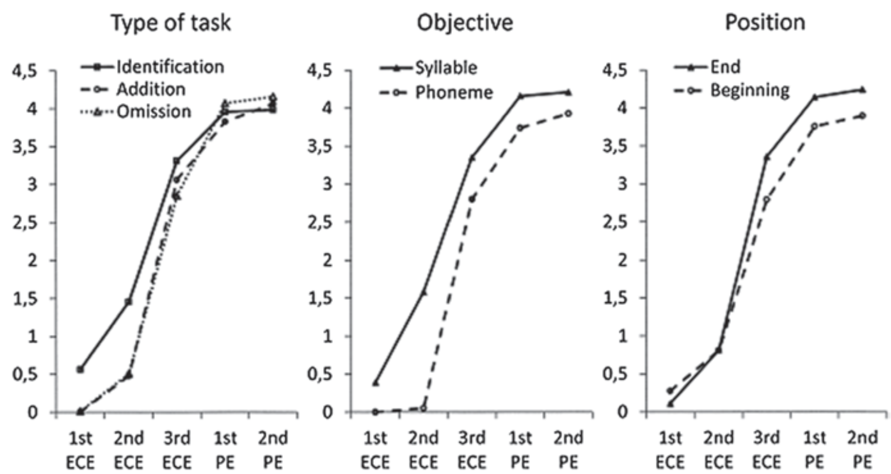

Figure 2. The Phonological Awareness Tasks' Main Effects.

As for the rhyme test, which is purely recognition-based (in the task, neither position nor demand gets manipulated), the only analysis possible was by year in school. In that respect, one-way analysis of variance revealed significant differences, and post-hoc tests yielded the same subgroups as those based on total scores, although the effect size was much smaller in this case: $F(1, 4) = 34.370$, $p < .001$, $\eta^2_p = .29$, $1–\beta = 1.0$.

Though difficulty is the most essential component of item analysis for any skills test (Wilson, 2005), item discrimination indexes provide important information by relating each item to the overall test. Corrected item-total correlations (where the item in question's score is not included in the total score) were computed to measure discriminant power, that is, the ability of each task to differentiate adequately among students with different levels of the construct that the test measures. According to Table 3, except in ECE first-years, practically all the tasks' indexes reflected high discriminant power (≥ .40). Phoneme Identification in ECE second-years and Rhyme Recognition in ECE third-years did not meet that standard, but had acceptable results nonetheless (.30 ≤ .374 ≤ .39). Conversely, syllable identification had little discriminant power in ECE second-years, and for ECE first-years, discriminant power was low across the board. However, that is not to say the items were not useful in other ways.

Obviously, by employing smaller, more homogenous samples like the subgroups created by year in school, reliability dropped. Results nonetheless indicated good (ECE second-years) or very good (all other years in school) reliability.

IRC Subscale Analysis

The relationship between total scores on this subscale and academic level appears in Table 4.

Here, too, analysis of variance revealed significant differences, both in number of correct answers, which increased as schooling progressed , $F(4, 336) = 197.897$, $p < .001$, and Welch $(5,107.275) = 410.550$, $p < .001$), and task completion time, which decreased as education progressed, $F(4, 335) = 47.048$, $p < .001$, and Welch $(4, 130.276) = 59.904$, $p < .001$. The effect size was larger in the first case than in the second ($\eta^2_p = .702$ and $\eta^2_p = .360$, respectively). In both cases, the power of statistical tests reached 1.0.

Table 3. *Discrimination Indices on the Test's Phonological Awareness Tasks*

Alpha =	1st ECE .14	2nd ECE .72	3rd ECE .84	1st PE .88	2nd PE .84
Rhyme Recognition	.061	.428	.366	.502	.483
Syllable Identification	.141	.259	.692	.536	.735
Phoneme Identification		.378	.651	.798	.756
Syllable Addition	−.114	.740	.655	.690	.720
Phoneme Addition			.667	.727	.807
Syllable Omission	.177	.728	.581	.666	.707
Phoneme Omission			.730	.764	.802

https://doi.org/10.1017/sjp.2015.15 Published online by Cambridge University Press

Table 4. *Average Proportion of Correct Responses and Time Used (Sec.s) on Each of the Test's Initial Reading Competence Tasks by Year in School*

Task		1st ECE	2nd ECE	3rd ECE	1st PE	2nd PE
		(*n* = 39)	(*n* = 50)	(*n* = 91)	(*n* = 74)	(*n* = 87)
Reading of:		*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
UcL	A	.38 (.25)	.78 (.17)	.81 (.22)	.94 (.14)	.95 (.15)
	S	115.24 (48.24)	79.07 (21.57)	52.29 (35.49)	38.20 (15.67)	34.42 (22.79)
LcL	A	.17 (.19)	.51 (.25)	.73 (.22)	.92 (.10)	.92 (.16)
	S	94.00 (42.86)	82.36 (22.55)	55.26 (52.03)	4 .65 (14.56)	36.13 (16.65)
RW	A			.65 (.37)	.94 (.18)	.94 (.19)
	S			38.16 (27.09)	17.75 (9.10)	15.46 (8.13)
PW	A			.50 (.35)	.92 (.20)	.93 (.20)
	S			42.27 (2 .16)	23.43 (1 .81)	2 .48 (8.84)
IW	A			.41 (.38)	.92 (.20)	.92 (.22)
	S			49.87 (23.26)	23.81 (11.48)	2 .04 (1 .43)
WS	A			.22 (.31)	.75 (.27)	.83 (.16)
	A			77.48 (39.81)	55.35 (28.54)	4 .16 (16.40)
Total	A	.16 (.13)	.39 (.14)	.66 (.24)	.92 (.13)	.93 (.16)
	S	106.13 (41.23)	77.87 (28.52)	65.50 (52.03)	34.58 (13.43)	3 .49 (18.31)

Note: UcL = Uppercase letters; LcL = Lowercase letters; RW = Simple words; PW = Pseudowords; IW = Complex words; WS= Word splitting; A (accuracy) = Proportion of correct responses; S (speed) = Response time, in seconds.

On the other hand, post hoc tests indicated the proportion of correct answers was about the same in the PE first- and second-year subsamples. The same was true for PA: all differences in means were found to be statistically significant where $p < .001$, except between PE first- and second-years (difference = –.004, $p = 1.00$). Likewise, ECE second- and third-years used about the same amount of time to complete tasks: difference between PE first- and second-years = 4.09, $p = .48$; difference between ECE second- and third-years = 12.37, $p = .36$.

Differences between tasks remained as school went on, but the magnitude of those differences decreased, and different profiles emerged in terms of correct responses and time used. Grouping tasks according to their required skills, three groups emerge (letter reading, word reading, and word separation). Of those, letter reading got the most correct answers, but participants responded fastest on word reading tasks. The word separation task, on the other hand, had the fewest correct answers and took the longest to solve, indicating it was the most difficult at every academic level. Furthermore, we observed that within these groups, response speed was generally slower on tasks requiring higher accuracy (Figure 3).

Separately analyzing the three task types, significant differences in accuracy as well as speed became apparent (except for letter reading time), but the differences had a very small effect size (Table 5).

With regard to difficulty indexes at different academic levels, Figure 4 shows that ECE first- and second-years only managed to solve letter identification tasks. The subgroups differed in that for first-years, uppercase letter identification was difficult, and lowercase letter identification was very difficult, while for second-years, the first was easy and the second had medium difficulty. At the other end of the spectrum were PE first- and second-years, for whom all the tasks were very easy, although the Word Separation task was borderline for that category in first-years. The most remarkable data occurred in the ECE third-year subsample, who showed a linear progression in the difficulty of different subtests, from very easy to very difficult.

In terms of different tasks' discriminant capacity (Table 6), they all showed high discriminant power (≥. 40) except Word Separation in ECE third-years. In that case, its discriminant power was acceptable (.30 ≤ .323 ≤ .39). Furthermore, reliability measures were very good in all subgroups.

Validity

To determine LolEva's construct validity, Exploratory Factor Analysis was conducted, excluding Confirmatory Factor Analysis for now because as other authors have pointed out (Pérez-Gil, Chacón, & Moreno, 2000), it would be redundant. Before proceeding with EFA, we confirmed that the following conditions were met:

https://doi.org/10.1017/sjp.2015.15 Published online by Cambridge University Press

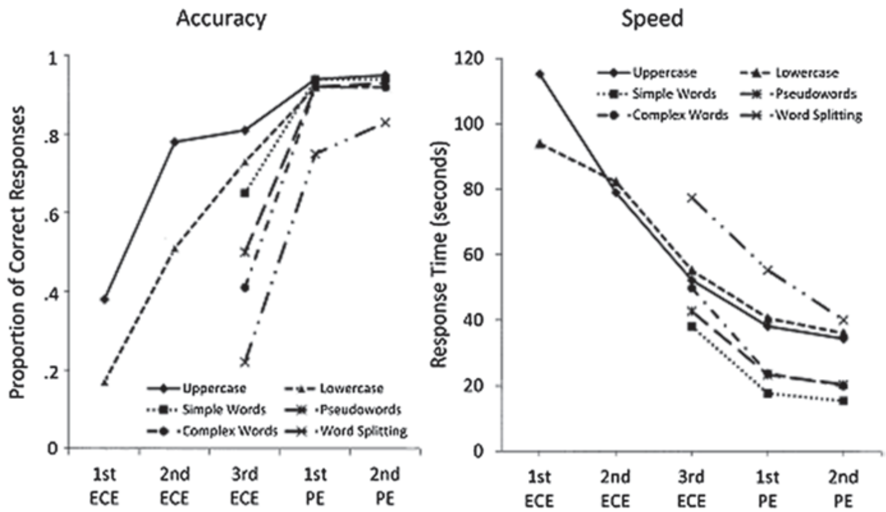

Figure 3. Speed and Accuracy Response Profiles on Initial Reading Competence Tasks.

more than half the correlations between variables were above .30, Bartlett's test of sphericity allowed us to reject the null hypothesis that the correlation matrix was an identity matrix, χ^2 (300) = 10371.73, $p < .001$, and the Kaiser-Meyer-Olkin measure of sampling adequacy (Kaiser, 1974) yielded a value that is considered optimal: *KMO* = .952.

A principal components extraction procedure was carried out, applying the criterion that components with initial eigenvalues > 1 get extracted. Equamax rotation (reducing the number of variables with high loadings on one factor, and the number of factors needed to explain one variable) was later applied. This yielded three factors, each explaining rather similar proportions of total variance, and together explaining 75.03%. The results appear in Table 7 (loadings below 0.30 were eliminated).

As the table illustrates, the first factor is basically made up of tasks from the PA subscale. Meanwhile, the IRC subscale comprises factors two and three, the latter including time measures on the reading and word splitting tasks. Upper and lowercase letter reading times are part of the second factor, but their factor loadings were negative.

Discussion

Our analyses have confirmed the test's high reliability, with alpha values of .94 for PA and .92 for IRC, and a factor structure that explains a high percentage of

Table 5. *Results of Repeated Measures Analyses by Initial Reading Competence Task Type*

Task	Measure	*df*	*F*	η^2_p	1-β [a]
Letter Reading	Accuracy	1	392.041***	.547	1.000
	Speed	1	.704	.002	.133
Letters x Year in School	Accuracy	4	50.210***	.382	1.000
	Speed	4	8.234***	.092	.999
Error (Letters)	Accuracy	325			
	Speed	325			
Word Reading	Accuracy	1	71.766***	.252	1.000
	Speed	1	28.860***	.119	1.000
Words x Year in School	Accuracy	2	31.463***	.228	1.000
	Speed	2	5.802***	.052	.867
Error (Words)	Accuracy	213			
	Speed	213			
Word Splitting (Accuracy/Speed)		1	716.501***	.790	1.000
Word Splitting x Year in School		2	19.556***	.171	1.000
Error (Word Splitting)		190			

[a]Calculated where alpha = .05
***$p < .001$

https://doi.org/10.1017/sjp.2015.15 Published online by Cambridge University Press

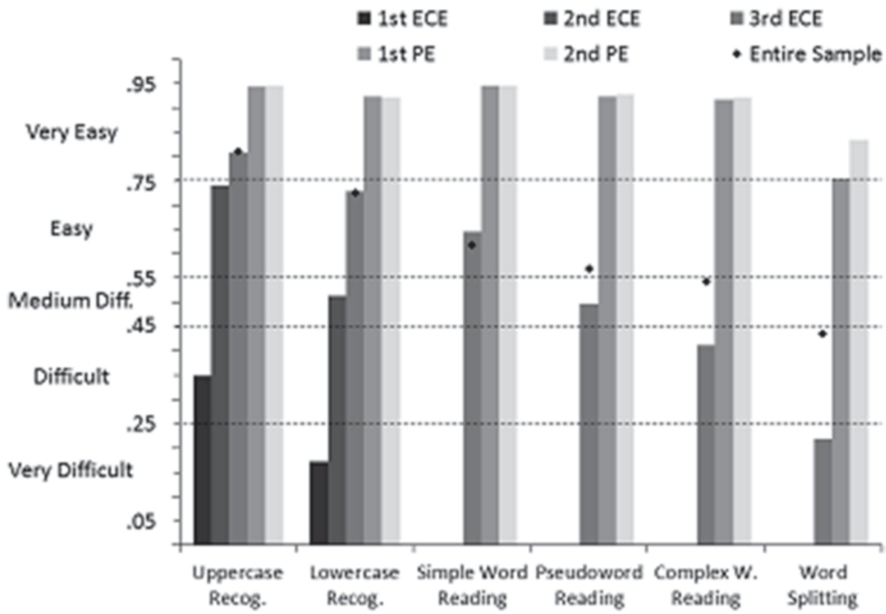

Figure 4. The Initial Reading Competence Tasks' Difficulty Indexes.

variance (75%). Furthermore, they show that the tasks utilized, and their respective items, have the psychometric properties needed to statistically guarantee the test and its usefulness as a tool in the early detection of learning difficulties. Scores on this test had comparable if not higher reliability than other similar tools with a pencil-and-paper format (the PROLEC-R by Cuetos et al., 2007; the TALE-2000 by Toro et al., 2000; and the PSL by Jiménez & Ortiz, 1995), apart from the age range covered.

On the other hand, results on the PA and IRC subscales (both in terms of accuracy and speed) suggested a distinct growth process in development between the first year of ECE and the beginning of PE. The way writing is introduced, socially and educationally, seems to influence that development.

In fact, while children in their first year of ECE become able to identify rhyme and recognize syllables, they can only identify uppercase (in the learning-to-read process, those are introduced first, so logically, their outcomes are best) and lowercase letters with difficulty. In the second year of ECE, they can to some extent complete PA tasks (except phoneme addition and omission). At that point they can also recognize upper and lowercase letters, like in the year before, but now with ease. Reading words and sentences is still absent at this stage. Thus, it is the level of difficulty posed by letter reading that differentiates these groups from one another.

In the third year of ECE, there is a qualitative and quantitative jump. Students are able to do all the PA tasks, with different levels of difficulty, and easily or very easily read uppercase letters, lowercase letters, and simple words. They can also read pseudowords of medium difficulty, but words with complex grapheme-phoneme agreement (referred as *complex* in this paper for the sake of brevity) and splitting sentences into words remain difficult and very difficult for them, respectively. Therefore, during this year in school, lexical access is already possible via both routes – direct or phonological – and they also display the ability to separate words at every level.

Regarding our results pertaining to phonological awareness, the developmental sequence we observed

Table 6. *Discrimination Indices on the Test's Initial Reading Competence Tasks*

Alpha =	1st ECE .88	2nd ECE .88	3rd ECE .91	1st PE .87	2nd PE .92
Uppercase Letter Reading	.825	.788	.820	.525	.913
Lowercase Letter Reading	.825	.788	.898	.844	.925
Simple Word Reading			.897	.860	.914
Complex Word Reading			.834	.854	.843
Pseudoword Reading			.935	.817	.887
Word Splitting			.323	.643	.707

Table 7. *Results of Factor Analysis on the LolEva Test with Equamax Rotation*

Components	I	II	III
% Explained Variance	29.49	29.13	16.41
Rhyme Recognition	**.805**		
First Syllable Identification	**.627**		
Last Syllable Identification	**.685**	.330	
Addition of First Syllable	**.694**	.461	
Addition of Last Syllable	**.590**	.487	.387
Omission of First Syllable	**.570**	**.628**	
Omission of Last Syllable	**.587**	.567	
Identification of First Phoneme	**.622**	.540	.327
Identification of Last Phoneme	**.682**	.419	.320
Addition of First Phoneme	**.751**		.359
Addition of Last Phoneme	**.664**	.470	.383
Omission of First Phoneme	**.711**	.496	
Omission of Last Phoneme	**.608**	.573	.323
Reading Uppercase Letters	.319	**.744**	
Uppercase Letter Reading Time		**−.818**	
Reading Lowercase Letters	.453	**.739**	
Lowercase Letter Reading Time		**−.785**	
Simple Word Reading	.457	**.774**	
Complex Word Reading	.566	**.726**	
Pseudoword Reading	.519	**.767**	
Word Separation	**.615**	**.575**	
Simple Word Reading Time			**.886**
Complex Word Reading Time			**.939**
Pseudoword Reading Time			**.936**
Word Splitting Time	.356	.362	**.439**

Note: Factor loadings < .30 were excluded.

coincides with the levels described in Treiman's (1991) hierarchical model; and the five levels of difficulty Adams (1990) identified on phonological awareness tasks according to linguistic, analytic, and memory demands; and the growing body of data on phonological awareness development (Jiménez & Ortiz, 1995). Conversely, these results only partially support those of Anthony et al. (2013). They were consistent with those authors' findings about the unidimensionality of PA, but not about what elements define task complexity and difficulty. Task complexity seems to play an important role in the development of phonological awareness skills. However, a word's structural complexity, the position and type of segment being manipulated within the word, and response format are all included under the concept of *complexity*. In this study, PA task complexity or difficulty referred to those two factors interacting in such a way as to impact the expression of a skill that seems to have just one factor. On that point, this study found that skills' different levels of difficulty matched their order of emergence.

Certain differences between our two studies make it challenging to compare them; task type was not entirely consistent. The present research utilized multiple-choice tasks, or similar, to assess rhyme recognition, beginning and ending syllable recognition, and beginning and ending phoneme recognition. In all other cases (omission and addition), free response tasks were employed. The fact that all PA tasks were grouped into a single factor may support the notion of one-dimensionality these authors suggest, but the influence of developmental sequence seems to include additional factors it interacts with reciprocally (working memory and executive functions, processing speed, bilingual context, etc.). The data presented here can be said to reflect that sort of interaction by showing that with age, free recall tasks have the same level of difficulty as recognition tasks (Identification is easier than Addition and Omission but their difficulty/ease tends to equalize as academic level increases); syllable-involved tasks are easier than phoneme-involved tasks at all academic levels; and finally, once respondents can manipulate syllables as well as phonemes (from the third year of ECE on), it is easier when the segment involved – syllable or phoneme – is at the end of the word than when it is at the beginning. Phonological awareness involves knowledge, capacity for manipulation, supervision, and conscious control (like other aspects of metalanguage), and requires interactive feedback from other top-down and bottom-up processes.

The results of this research, therefore, seem to indicate that although phonological awareness has just one factor, it follows one course or another as a function of other variables (cognitive and contextual) that, while independent, favor or hinder its development with age.

ECE first- and second-year students' difficulty manipulating the phoneme segments of words, and third-years' ease in doing so, appear to be the result of not only their levels of metaphonological development, but also their capacity to manage other cognitive-type functions (like executive functions and naming speed) that, while independent of PA, have a well established effect on reading acquisition (Caravolas et al., 2013; Mayor et al., 2012; Peralbo et al., 2012). This situation, paired with increasing, more systematic exposure to reading in both school and extracurricular contexts, may explain how ECE third-years could already exhibit the hothouse effect posited by Torgesen and Davis (1996), and Defior (2008), who suggested that "phonological awareness may act like an enzyme that helps establish a more comfortable context for learning written language" (p. 344).

In the present research, the difficulty of reading processes increased much more linearly: uppercase letters, lowercase letters, simple words, pseudowords, complex words, and splitting sentences into words. Furthermore, the more correct responses a student gave, the less reading time they needed. That is essentially consistent with what we know about reading acquisition and the processes it involves (Cuetos, 2008).

This study had two main limitations. The first has to do with the number of participants in the ECE first-year group. That subsample's size should be increased in order to corroborate these results which, while theoretically consistent, need more complete empirical validation. The second limitation would also require a considerably larger sample size. Given our interest in determining the role of variables related to social-economic-cultural context in the level and progression of metalinguistic development (family's level of education, monolingual vs. bilingual context, public or private school, etc.), taking such contextual variables into consideration would help us better understand not only *what* develops (in which case the psycholinguistic perspective has been predominant), but *how* it develops (basic knowledge for implementing any educational or clinical approach).

Subsequent studies will allow us to confirm LolEva's underlying factor structure through Confirmatory Factor Analysis, and to ascertain its concurrent, discriminant, and predictive validity. To do so will require, on the one hand, analyses to examine its relationship with other phonological awareness scales (like the ones highlighted in this paper), and on the other, analyses to identify the profile associated with atypical forms of development, whether the result of learning difficulties or more or less severe developmental disorders.

References

Adams M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: Bolt.

Alegría J. (2006). Por un enfoque psicolingüístico del aprendizaje de la lectura y sus dificultades –20 años después [Support for a psycholinguistic approach to reading acquisition and reading difficulties – 20 years later]. *Infancia y Aprendizaje*, *29*, 93–111. http://dx.doi.org/10.1174/021037006775380957

Anthony J. L., Williams J. M., Durán L. K., Gillam S. L., Liang L., Aghara R., ... Landry S. H. (2011). Spanish phonological awareness: Dimensionality and sequence of development during the preschool and kindergarten years. *Journal of Educational Psychology*, *103*, 857–876. http://dx.doi.org/10.1037/a0025024

Brady S. A., & Shankweiler D. R. (Eds.) (1991). *Phonological processes in literacy*. Hillsdale, NJ: Erlbaum.

Bryant P., & Bradley L. (1985). *Childhood reading problems*. Oxford, UK: Blackwell.

Cardinal R. N., & Aitken M. R. F. (2006). *ANOVA for the behavioural sciences researcher*. New Jersey, NJ: Erlbaum.

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

Carrillo M. S., & Marín J. (1996). *Desarrollo metafonológico y adquisición de la lectura: Un programa de entrenamiento* [Metaphonological development and reading acquisition: A training program]. Madrid, Spain: CIDE.

Carroll J. M., Snowling M. J., Stevenson J., & Hulme C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, *39*, 913–923. http://dx.doi.org/10.1037/0012-1649.39.5.913

Cuetos F. (2008). *Psicología de la lectura* [Psychology of reading] (7th Ed.). Madrid, Spain: Kluwer.

Cuetos F., Rodríguez B., Ruano E., & Arribas D. (2007). *PROLEC-R: Batería de evaluación de los procesos lectores -revisada* [Test battery for reading processes assessment -revised]. Madrid, Spain: TEA.

Defior S. (1996). Una clasificación de las tareas utilizadas en la evaluación de las habilidades fonológicas y algunas ideas para su mejora [A classification of tasks utilized in phonological skills assessment, and some ideas on how to improve them]. *Infancia y Aprendizaje*, *19*, 49–63. http://dx.doi.org/10.1174/02103709660560546

Defior S. (2008). ¿Cómo facilitar el aprendizaje inicial de la lectoescritura? Papel de las habilidades fonológicas [How to facilitate initial literacy acquisition? The role of phonological skills]. *Infancia y Aprendizaje*, *31*, 333–345. http://dx.doi.org/doi:10.1174/021037008785702983

Gillam R., & van Kleeck A. (1996). Phonological awareness training and short-term working memory: Clinical implications. *Topics in Language Disorders*, *17*, 72–81. http://dx.doi.org/10.1097/00011363-199611000-00008

Howell D. C. (2013). *Statistical methods for psychology* (8th Ed.). Belmont, CA: Wadsworth.

Jiménez J. E., & Ortiz M. R. (1995). *Conciencia fonológica y aprendizaje de la lectura: Teoría, evaluación e intervención* [Phonological awareness and reading acquisition: Theory, assessment and intervention]. Madrid, Spain: Síntesis.

Jiménez J. E., Antón L., Díaz A., Guzmán M. R., Hernández-Valle M. I., Ortiz M. R., … Muñetón M. A. (2007). *SICOLE-R-Primaria: Manual de uso e instrucciones para el examinador* [SICOLE-R-Basic: Manual and examiner instructions]. Santa Cruz de Tenerife, Spain: Ocide.

Kaiser H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 31–36. http://dx.doi.org/10.1007/BF02291575

Kirk R. E. (2013). *Experimental design: Procedures for the behavioral sciences* (4th Ed.). Thousand Oaks, CA: Sage.

Márquez J., & de la Osa P. (2003). Evaluación de la conciencia fonológica en el inicio lector [Assessing phonological awareness in initial reading]. *Anuario de Psicología*, *34*, 357–370.

Leong C. K. (1991). From phonemic awareness to phonological processing to language access in children developing reading proficiency. In D. J. Sawyer & B. J. Fox (Eds.), *Phonological awareness in reading: The evolution of current perspectives* (pp. 217–254). New York, NY: Springer-Verlag.

Mayor M. A., Fernández M. L., Tuñas A., Zubiauz B., & Durán M. (2012). La relación entre funciones ejecutivas y conciencia fonológica en Educación Primaria [The relationship between executive functions and phonological awareness in Primary School]. In L. Mata, F. Peixoto, J. Morgado, J. Castro, & V. Monteiro (Eds.), *Educaçao, aprendizajem e desenvolvimento* [Education, learning and development] (pp. 1792–1806). Lisbon, Portugal: ISPA.

Morais J. (1991). Phonological awareness: A bridge between language and literacy. In D. J. Sawyer & B. J. Fox (Eds.), *Phonological awareness in reading. The evolution of current perspectives*. Springer Series in Language and Communication (Vol. 28, pp. 31–71). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4612-3010-6_2

Muñiz J. (1999). *Psicometría* [Psychometrics]. Madrid, Spain: Pirámide.

Peralbo M., Brenlla J. C., García M., Barca A., & Mayor M. A. (2012). Las funciones ejecutivas y su valor predictivo sobre el aprendizaje inicial de la lectura en educación primaria [The executive functions and their predictive value in initial reading acquisition during Primary School]. In L. Mata, F. Peixoto, J. Morgado, J. Castro, & V. Monteiro (Eds.), *Educaçao, aprendizajem e desenvolvimento* [Education, learning and development] (pp.76–90). Lisbon, Portugal: ISPA.

Pérez-Gil J. A., Chacón S., & Moreno R. (2000). Validez de constructo: El uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez [Construct validity: The use of exploratory-confirmatory factor analysis to collect evidence of validity]. *Psicothema*, *12*, 442–440.

Stanovich K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York, NY: Guilford Press.

Torgesen J. K., & Davis C. (1996). Individual difference variables that predict response to training in phonological awareness. *Journal of Experimental Child Psychology*, *63*, 1–21. http://dx.doi.org/10.1006/jecp.1996.0040

Toro J., Cervera M., & Urio C. (2000). *EMLE-TALE 2000. Escalas Magallanes de Lectura y Escritura* [Magallanes Reading and Writing Scales]. Baracaldo, Spain: Albor-COHS.

Treiman R. (1991). Phonological awareness and its roles in learning to read and spell. In D. J. Sawyer & B. J. Fox (Eds.), *Phonological awareness in reading. The evolution of current perspective* (pp. 159–189). New York, NY: Spinger-Verlag.

Wilson M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum.

Zimmerman D. W. (2004). A note on preliminary tests of equality of variances. *British Journal of Mathematical and Statistical Psychology*, *57*, 173–181. http://dx.doi.org/10.1348/000711004849222