

Analysis of performance and improvement in Reading and Writing in Preschool

CITA:

González-Valenzuela, M.J., **Martín-Ruiz, I.**, Prieto, G., & Rivas-Moya, M.T. (2018). Analysis of performance and improvement in Reading and Writing in Preschool, *Revista de Educación*, 382, 225-247.

<http://dx.doi.org/10.4438/1988-592X-RE-2018-382-398>

Abstract

This study analyzes the performance and improvement in reading and writing in the last two years of preschool, according to age and sex. The psychometric properties of the tests used to evaluate these skills are also examined. One hundred and thirty 4-year-old children and 144 five year olds (44.6% and 45.1% males and 55.4% and 54.9% females, respectively), attending public (52.2%) and semi-private schools (47.8%) and without physical or psychological disabilities, were evaluated using a reading and a writing test which integrated performance subtests: reading accuracy, reading comprehension, copying and transfer of information and dictation. Two assessment measures were carried out (pre-test and post-test) and most of the testlets resulted in pre- and post-test scores that were well differentiated and reliable. Differences were found with respect to age but not sex. In addition, there was an improvement in the post-test scores of both ages that was similar for both boys and girls. Our findings support the usefulness of instruments created to measure performance in reading and writing at these ages, as well as the promotion of reading and writing from an early age as a means for maintaining these results later on.

Key Words: preschool, reading, writing, age, sex, reliable change

Introduction

There are very few studies that evaluate the performance and improvement in reading and writing at an early age, due, in part, to the debate in the scientific and academic communities on when the teaching of reading and writing should begin.

One of the standpoints regarding the right moment for which this should commence claims that learning to read and write should be subject to individual maturity. Therefore, it makes sense that this would begin from the age of six when children are prepared and their psychological maturity allows for adequate learning (Revuelta & Guillen, 1987). According to these ideas, individual maturity marks the onset of learning. This position presents the disadvantage that sometimes the opportunity is lost to exploit the high learning potential that children have at an early age (González & Delgado, 2006, González-Valenzuela & Martín-Ruiz, 2017).

From other theoretical perspectives, however, it can be argued that the start of the teaching-learning process of written language should not be exclusively subject to the maturity of the individual, but should rely more on early intervention that would help to develop maturity. Some of these studies defend that learning literacy skills should begin as early as three and four years of age (Brand & Dalton, 2012; Brown, Scull, Nolan, Raban, & Deans, 2012; Dunphy, 2012; Elliott & Olliff, 2009; Slavin, Madden, Dolan, & Wasik, 1996; Vadasy & Sanders, 2008; Wright, Stackhoyse, & Wood, 2008). In addition, many of these studies find that learning to read and write from a young age improves a child's academic performance, not only with respect to written language but also in other areas such as mathematics. It also reduces the number of children at risk of failure and those with specific learning difficulties (González & Delgado, 2006, 2009; González-Valenzuela & Martín-Ruiz, 2017; González, Martín, & Delgado, 2011, 2012; Lovett, Frijters, Wolf, Steinbach, Sevcik, & Morris, 2017, Slavin & Madden, 2001, Suggate, Schaughency, & Reese, 2013, Park, Chaparro, Preciado, & Cummings, 2015). This study is based on the idea that children can and should learn from an early age, and that also evident changes can be seen in their learning during this period according to age. These positions about when to start the teaching-learning process of written language are linked to the legislation of different educational standards. Thus, there are countries where learning to read and write starts from as early as three to four years and others, which are the vast majority, which initiate the process from the age of six. Owing to this situation, the creation of tests evaluating reading and writing performance at an early age is not common practice. Most of the tests generated are taken after the age of 6, when legislation stipulates in most countries that learning to read and write is compulsory.

There are also few studies that evaluate the change in performance in reading and writing at early ages according to sex, despite the fact that there are data indicating differences between boys and girls with respect to written language in primary and secondary education.

The differences in terms of sex in relation to written language are periodically analyzed in different international studies. The Progress in International Reading Literacy Study (PIRLS) evaluates fourth grade students, and the reports from 2006 and 2017 highlight that girls perform better in reading than boys in all of the participating countries (MECD, 2017, Mullis, Martin, González, & Kennedy, 2003; Mullis, Martin, Kennedy, & Foy, 2007). In addition, the Programme for International Student Assessment (PISA), which evaluates students during the last years of secondary education, also shows differences that favor girls regarding reading in all of the OECD countries and the EU, as well as in Spain (MECD, 2016; OECD, 2010a). It has been observed that the gap between boys and girls in reading has increased during 2000 to 2012 (Gomendio, 2015). Martínez García & Córdoba (2012) also find differences in favor of girls and add that they are associated with the mother's economic activity and her level of studies, affecting the girls more positively than the boys. Pargulski & Reynolds (2017) find differences again in favor of girls in reading and writing, specifically in reading fluency, essay writing and sentence composition due to intrinsic motivation; these findings are in line with those of other researchers (Reynolds, Scheiber, Hajovsky, Schwartz, & Kaufman, 2015; Schwabe, McElvany, & Trendtel, 2014). Additional studies point to neurological differences, in particular, to differences in brain development, which are amplified or reduced by sociocultural contexts. This would explain the differences between countries in terms of inequality between boys and girls in reading (Wood & Eagly, 2002).

Furthermore, it seems that differences in terms of sex are on the increase from primary to secondary education. This may be due to the discrepancies in reading motivation manifested by both groups, where motivation is seen to be greater in girls (Martínez García & Córdoba, 2012; Logan & Johnston, 2009; OECD, 2010b). Logan & Johnston (2009) found that girls have better scores than boys in reading comprehension, reading activity and attitudes toward reading; however, the differences in reading comprehension were minor. Along these same lines, a study conducted by Troia, Harbaugh, Shankland, Wolbers, & Lawrence (2013) on adolescents found that age, sex, and motivation toward writing influenced the quality of the students' essays, being more effective in girls. Camarata & Woodcock (2006) also find that from preschool to primary education, girls score better in reading and writing fluency, but not in intellectual capacity.

In this context, the main objective of our study was to analyze the performance and improvement in reading and writing of 4- and 5-year-old children in preschool education, according to age and sex, by analyzing the psychometric properties of tests designed to evaluate literacy skills at these specific ages.

Method

Participants

The sample population was comprised of 274 schoolchildren in the second and third year of preschool education (4- and 5-year-olds), attending public and semi-private schools and with no physical or psychological disabilities.

The schools, located in the province of Malaga (Spain), had been classified based on their socio-cultural status (high, medium or low). According to a school census published by the Ministry of Education of the Regional Government of Andalusia, twelve public and semi-private primary schools were considered to have a medium socio-cultural status. Of these, five were chosen randomly to participate in the study. In this area, 30% of the population was below the average level of illiteracy (González & Delgado, 2009). Among the 274 participants, 143 (52.2%) attended public schools and 131 (47.8%) semi-private schools. According to the psycho-pedagogical reports provided by the school counselors, none of the participants presented physical or psychological disabilities.

A total of 130 four-year-olds ($M = 4.4$ and $SD = .5$) took the reading and writing pre-test and 114 took the corresponding post-test. Also, 144 five-year-olds ($M = 5.3$ and $SD = .5$) took the reading and writing pre-test and 135 children took the corresponding post-test.

Table I shows the number and percentage of school children, according to sex and age, who participated in taking the pre- and post-tests.

TABLE I. Distribution of children, according to sex and age, who took the pre- and post-tests

Age (years)	Sex					
	Pre-test			Post-test		
	Boys	Girls	Total	Boys	Girls	Total
4	58 (44.6%)	72 (55.4%)	130	53 (40.8%)	61 (46.9%)	114
5	65 (45.1%)	79 (54.9%)	144	62 (43.1%)	73 (50.7%)	135

Total	123	151	274	115	134	249
-------	-----	-----	-----	-----	-----	-----

Instruments

Reading was evaluated using a test that integrated 12 testlets from the Performance in Reading Accuracy (REL) and Performance in Reading Comprehension (RCL) tests (González & Delgado, 2006, González et al., 2011, González-Valenzuela & Martín-Ruiz, 2017). REL evaluates the answers obtained for the identification of letters, the identification of syllables of different linguistic structure (CV, CV, CVC, CCV, CCVCC), the identification of words of different length and frequency, the identification of pseudowords of different lengths, the identification of sentences with different lengths and complexity and the identification of texts of different lengths, familiarity of content and grammatical complexity. RCL evaluates the comprehension of phrase words (complete sentences) and texts.

Each testlet is composed of items referring to the same content and dichotomous response (correct or incorrect). The score in each testlet is the sum of the correct answers to the items. The testlets that makes up the reading test and the number of items in each testlet are: reading letters (Letule, 28), syllables (Letusi, 16) words (Letupa, 16) pseudowords (Letupe, 8), sentences (Letufra, 6), texts (Letute1 to Letute4, 12, 19, 47, 55, respectively); and the understanding of words (Compa, 14), sentences (Comfra, 6) and texts (Comte, 18). The total score of the reading test (256 items) is the sum of the testlet scores.

Writing was evaluated by integrating 7 testlets from the Performance in Copy (REC) and Performance in Dictation (RED) tests (González & Delgado, 2006, González et al., 2011). REC evaluates the responses to copying letters, copying words of different length and frequency and copying sentences of different length and linguistic structure. The RED test evaluates the responses to the dictation of letters, dictation of words of different length, structure and frequency, dictation of pseudowords of different length and structure, dictation of sentences and dictation of texts of different length and linguistic structure.

The testlets that make up the writing test and the number of items in each testlet are: copying letters (Copile, 7), words (Copipa, 8), phrases (Copifra, 8); and the dictation of letters (Ditale, 8), words (Ditapa, 8), pseudowords (Ditapse, 6) and sentences (Ditafra, 16). The total score of the writing test (61) is the sum of the testlet scores.

Procedure

The design was quasi-experimental and included two evaluation measures (pre-test and post-test): the pre-test was taken at the beginning of the academic year in October; and the post-test was taken when the students had finished the academic year in May. Both evaluations were carried out on each group of students in the second and third year of preschool education.

After permission had been provide by the schools and the parents had signed the informed consent forms, the test was administered individually to each student by Psychology graduates.

The test took approximately thirty minutes for each child to complete. The reading tests were administered first, followed by the writing tests.

Data Analysis

Initially, a classic analysis of the psychometric characteristics of the reading and writing tests was carried out.

In the case of the testlets, the mean, standard deviation, the corrected index of discrimination and the percentage of the average with respect to the highest possible score were calculated. The corrected index of discrimination is the correlation between the testlet and the test scores after subtracting the testlet ($R_t(X-t)$). **An indicator of the mean level in the variable is the ratio between the mean and the possible maximum score (multiplied by one hundred).**

: values higher or lower than 50% indicate a high or low performance, respectively.

The mean, standard deviation, maximum score, minimum score, percentage of the mean with respect to the highest possible score and the internal consistency reliability (Beta coefficient of Raju, 1977) were calculated for both the pre-test (identified with the number 1) and post-test (identified with the number 2). In addition, statistical contrasts were conducted in order to analyze for differences in performance according to sex and age.

The analysis for identifying improvements in the pre- and post-test scores was carried out using both statistical and psychometric approaches. The statistical approach focused on analyzing the size and significance of the difference between the means of the pre- and post-test scores in general and according to sex and age. To this end, **in independent samples, the Welch's t-test was used to test hypotheses for means difference, and effect size (d) (Cohen, 1988) was calculated in order to quantify the size of the means difference.**

In the psychometric approach, an individual indicator of reliable improvement (not attributable to measurement errors) was used as proposed by Speer (1992). In this way, the improvement is considered reliable when the score X_2 is greater than the cutoff point (PC) corresponding to the upper limit of the estimation interval of the true pre-test score (confidence level).

$$PC = [r_{xx}(X_1 - M_1) + M_1] + 2 S_1 (1 - r_{xx})^{1/2}$$

where, r_{xx} is the reliability coefficient in the pre-test (Beta1); X_1 is the individual score in the pre-test; and M_1 and S_1 are the mean and standard deviation of the scores in the pre-test, respectively.

The cutoff point allows the schoolchildren to be classified in one of two categories (no improvement or a reliable improvement). This manner of classification facilitates the calculation of the percentage of students that show a reliable improvement in their post-test scores. Reliable improvement was analyzed in general and according to sex and age. The test χ_1^2 , applied to the 2x2 contingency tables, is used to test the hypothesis regarding the difference in the proportion of students that show a reliable improvement in the post-test, taking into account the variables of sex and age. The size of the degree of association (Cohen, 1988) between the reliable improvement in relation to reading or writing and sex/age is given.

The size of the d effect was calculated using the G * Power 3 program (Faul, Erdfelder, Lang & Buchner, 2007), and the data analysis was carried out using the Statistical Package for the Social Sciences (SPSS, 2006) version 14.0.

Results

Performance and reliability of the reading and writing tests

Table II shows the testlets included in the pre- and post-tests and the statistics corresponding to the reading test.

In the pre-test, performance on average was low (the mean is only 2.8% of the highest possible score) and variability was low in both the test and in all of the testlets that comprised it.

The post-test, however, showed an increase in average performance in the test (24.3% of the highest possible score) as well as in the testlets. In both cases, variability also increased.

TABLE II.
Statistics of the the Reading test items

<i>Name</i>	<i>M₁</i>	<i>S₁</i>	<i>X_{ma}_{x1}</i>	<i>Rt(X-t)₁</i>	<i>M₂</i>	<i>S₂</i>	<i>X_{max}₂</i>	<i>Rt(X-t)₂</i>
Letule	4.25 (15.2%)	5.05	28	.59	14.61 (52.2%)	9.18	28	.75
Letusi	0.19 (1.2%)	1.17	12	.86	3.83 (23.9%)	5.34	16	.93
Letupa	1.51 (9.4%)	2.81	12	.55	7.47 (46.7%)	5.33	16	.82
Letupse	0.07 (.9%)	0.60	6	.82	2.33 (29.1%)	3.19	8	.88
Letufra	0.05 (.8%)	0.44	5	.82	1.42 (23.7%)	2.32	6	.94
Letule1	0.08 (.7%)	0.85	11	.74	2.68 (22.3%)	4.90	12	.93
Letule2	0.15 (.8%)	1.43	18	.76	4.02 (21.2%)	7.59	19	.93
Letute3	0.06 (.1%)	0.97	16	.42	8.68 (18.5%)	17.87	47	.91
Letute4	0.02 (.0%)	0.30	5	.06	7.39 (13.4%)	18.59	55	.78
Compa	0.42 (3.0%)	1.51	13	.75	4.54 (32.4%)	6.05	14	.86
Comfra	0.05 (.8%)	0.42	4	.83	1.23 (20.5%)	2.23	6	.94
Comte	0.02 (.1%)	0.27	4	.65	1.49 (8.3%)	3.20	12	.91
Reading	6.87 (2.8%)	12.11	107	--	59.65 (24.3%)	76.93	239	--

Note--: *M₁* and *M₂*: Means and their percentages as a proportion of the highest score possible (%) on the pre- and post-tests, respectively; *S₁* and *S₂*: Standard deviations in the pre- post-test respectively; *X_{max1}* *X_{max2}*: Highest scores in the pre- and post-tests, respectively; *Rt(X-t)₁*: Task-test correct correlation (Reading) of the pre-test; *Rt(X-t)₂*: Task-test correct correlation (Reading) of the post-test; Source: The authors

The psychometric properties were considered acceptable, as there was a high level of discrimination among the testlets in the pre- (an exception is the low discrimination of the Letute testlet4) and post-tests. In addition, the reliability (0.83) of the test scores was good in the case of the pre-test and (0.97) excellent in the case of the post-test (Hernández, Ponsoda, Muñiz, Prieto, & Elosua, 2016).

Table III shows the testlets included in the pre- and post-tests and the statistics corresponding to the writing test.

In the pre-test, a low average performance is observed (the average in the test is only 9.2% of the highest possible score) as well as low variability. In addition, performance in 4 of the 7 testlets was dramatically low (the averages do not reach 1% of the highest possible score and the standard deviations had very low values). Moreover, most testlets showed small or very small discrimination indices. Low variability and the lack of discrimination among the several testlets can explain why the reliability (0.65) of the scores of the subtest in the pre-test can be considered adequate, but with some deficiencies (Hernández et al., 2016). The average performance, variability and reliability of the scores markedly improved in the post-test: 36.1% of the test takers achieved the highest possible score, standard deviation tripled and the reliability coefficient (0.93) was considered to be excellent.

TABLE III.
Statistics of the the Reading test items

<i>Name</i>	<i>M₁</i>	<i>S₁</i>	<i>X_{max1}</i>	<i>Rt(X-t)₁</i>	<i>M₂</i>	<i>S₂</i>	<i>X_{max2}</i>	<i>Rt(X-t)₂</i>
<i>Copile</i>	3.61 (51.6%)	2.68	7	.67	5.97 (85.3%)	1.65	7	.51
<i>Copipa</i>	0.89 (11.1%)	1.50	4	.69	4.47 (55.9%)	3.01	8	.74
<i>Copifra</i>	0.01 (.1%)	0.18	3	.03	3.07 (38.4%)	3.13	8	.80
<i>Ditale</i>	1.08 (13.5%)	1.66	7	.70	4.06 (50.7%)	2.70	8	.82
<i>Ditapa</i>	0.02 (0.2%)	0.20	3	.29	1.41 (17.6%)	2.22	8	.85
<i>Ditapse</i>	0.01 (0.2%)	0.19	3	.25	1.10 (18.3%)	1.85	6	.83
<i>Ditafra</i>	0.02 (0.1%)	0.30	5	.19	1.96 (12.25%)	3.96	16	.74
<i>Writing</i>	5.64 (9.2%)	5.25	27	--	22.03 (36.1%)	15.38	61	--

Note--: M₁ and M₂: Means and their percentages as a proportion of the highest score possible (%) in the pre- and post-tests, respectively; DT₁ y DT₂: Standard deviations in the pre- post-test respectively; X_{max1} y X_{max2}: Highest scores in the pre- and post-tests, respectively; Rt(X-t)₁: Task-test correct correlation (Reading) of the pre-test; Rt(X-t)₂: Task-test correct correlation (Reading) of the post-test;

Source: The authors

Differences in performance associated with age and sex

Table IV shows the means and standard deviations of the pre- and post-test scores in reading and writing corresponding to the children's age and sex.

As expected, the average reading performance of the 5-year-old participants was clearly higher than that of the 4 year olds. The variability of reading performance was higher at 5 than at 4 years: the results of the Levene test showed that there were statistically significant differences among the variances in the pre-test ($F(1,170.98)=18.06$; $p<.05$) and the post-test, ($F(1,145.11) = 365.31$; $p<.05$).

The Student t-test, for independent samples, showed that the differences among the means in reading for both ages were statistically significant in the pre-test ($t_{170.98} = -5.59$; $p<.05$) and in the post-test ($t_{141.11} = -10.55$; $p<.05$).

The size of the d effect was medium (0.65) in the pre-test and large (1.25) in the post-test.

The average performance and variability of the boys' reading scores were similar to the girls' scores (Table IV). The results of the Levene test showed that there were no statistically significant differences among the variances in the pre- ($F(1,272) = 3.59$; $p>.05$) and post-tests ($F(1,246) = .13$; $p>.05$).

The Student t-test, for independent samples, showed that there were no statistically significant differences among the means of the pre- ($t_{272} = 1.65$; $p>.05$) and post-tests ($t_{247} = -.07$; $p>.05$). The size of the d effect was small in both the pre-test (0.21) and the post-test (0.008).

As in the case of reading, the average performance of the 5-year-old students in writing was higher than that of the 4-year-old students. Also, the variability of the performance in writing was greater at 5 than at 4 years: the results of the Levene test showed that there were statistically significant differences among the variances in the pre- ($F(1,215.07) = 53.98$; $p<.05$) and post-tests ($F(1,209.14) = 68.24$; $p<.05$).

The Student t-test, for independent samples, showed that the differences were statistically significant in the pre-test ($t_{215.07} = -15.21$; $p<.05$) and in the post-test ($t_{209.14} = -12.54$; $p<.05$). In addition, the sizes of the d effect were large (1.79 and 1.52 respectively).

The average performance and variability of boys' writing scores were similar to the girls' (Table IV). The results of the Levene test showed that there were no statistically significant differences among the variances ($F(1,272) = 2.14$; $p>.05$) in the pre-test and in the post-test ($F(1,247) = .23$; $p>.05$).

The Student t-test, for independent samples, showed that there were no statistically significant differences among the means of the boys' and girls' writing scores in the pre- ($t_{272} = -1.08$; $p>.05$) and post-tests ($t_{247} = -1.53$; $p>.05$). The size of the d effect was small in both the pre- (0.13) and post-tests (0.19).

TABLE IV. Means and standard deviations in reading and writing associated with age and sex

Test	M ₄	M ₅	S ₄	S ₅	M _{boys}	M _{girls}	S _{boys}	S _{girls}
Pre-reading	2.94	10.42	4.59	15.30	8.26	5.74	14.54	9.59
Pre-writing	1.96	8.96	2.44	4.88	5.26	5.95	5.22	5.27
Post-reading	15.59	97.14	17.28	87.51	59.30	59.95	79.05	75.40

Change between the pre- and post-test: improvements in reading and writing

Tables II and III show that in both reading and writing the means of the post-test are higher than those of the pre-test for the total sample of participants.

The Student t-test, for dependent samples, indicated that the differences among the means of the pre-test and post-test were statistically significant in reading ($t_{247} = 11.36$; $p < .05$) and in writing ($t_{248} = 21.74$, $p < .05$). The size of the effect was medium in reading ($d = 0.72$) and large in writing ($d = 1.38$).

The data appearing in Table IV allowed us to verify the improvement in reading and writing associated with the participants' age and sex, where the means of both variables are greater in the post-test than in the pretest for all of the ages and sex subgroups.

In reading, the increase in the post-test was large and statistically significant at 4 ($t_{113} = 8.98$; $p < .05$; $d = .84$) and at 5 years ($t_{133} = 11.88$; $p < .05$; $d = 1.03$). In writing the increase in the post-test was also large and statistically significant at 4 ($t_{113} = 14.84$; $p < .05$; $d = 1.29$) and at 5 years ($t_{134} = 20.31$; $p < .05$; $d = 1.75$).

Taking into account the improvement associated with sex, it was observed that in reading the increase in the post-test was of medium size and statistically significant in boys ($t_{112} = 7.44$; $p < .05$; $d = .70$) and in girls ($t_{134} = 8.55$, $p < .05$; $d = .74$). In writing the increase in the post-test was also large and statistically significant in boys ($t_{113} = 13.61$, $p < .05$, $d = 1.28$) and in girls ($t_{134} = 17.08$, $p < .05$, $d = 1.47$).

In addition to quantifying the differences between the means of the pre- and post-tests, the improvement in reading and writing was analyzed using a psychometric approach, where the number of participants who reliably improved in the post-test was counted.

Table V shows that of the 164 participants who reliably improved in reading (66.1%), the number of 5 year olds (81.3%) was greater than the number of 4 year olds (48.2%). The percentage difference was statistically significant ($\chi^2 = 30.12$, $p < .05$) and the degree of association between the improvement in reading and age was moderate ($\phi = .35$). With regards to sex, Table VI shows that the percentage of students with an improved reading performance was high and similar for both boys (65.5%) and girls (66.7%). The percentage difference was not statistically significant ($\chi^2 = .04$; $p < .05$) and the degree of association between the improvement in reading and sex was very low ($\phi = .01$).

Table V shows that of the 188 participants who improved reliably in Writing (75.5%), the number of those who improved in the 5-year sample (88.1%) was greater than in the 4-year sample (60.5%). The percentage difference was statistically significant ($p < .05$) and the degree of association between the improvement in reading and age was moderate ($\phi = .01$).

TABLE V. Improvement and age. Number and percentage of cases that reliably improved in the post-test

Variable	N	4 years (114)	5 years (135)	χ^2	Sig.	ϕ
----------	---	---------------	---------------	----------	------	--------

Reading	248	55 (48.2%)	109 (81.3%)	30.12	.00	.35
Writing	249	69 (60.5%)	119 (88.1%)	25.49	.00	.32

Source: The authors

Regarding sex, Table VI shows that the percentage of students who improved their performance in writing was high and similar in both boys (71.9%) and girls (78.5%). The percentage difference was not statistically significant ($\chi^2_1 = 1.45$; $p < .05$) and the degree of association between the improvement in writing and sex was very low ($\phi = .08$).

TABLE VI. Improvement and sex. Number and percentage of cases that reliably improved in the post-test

Variable	N	Boys (115)	Girls (134)	χ^2_1	Sig.	ϕ
Reading	248	74 (65.5%)	90 (66.7%)	0.04	.84	.01
Writing	249	82 (71.9%)	106 (78.5%)	1.45	.23	.08

Source: The authors

As an indicator of the average level of performance, the percentage corresponding to the average was used with respect to the highest possible score in reading and writing. It was considered that the improvement was reliable (not attributable to measurement errors) when the score in the post-test in reading or writing was greater than the cutoff point proposed by Speer (1992). The internal consistency in each test was calculated using the beta coefficient of Raju (1977).

The results indicated that in the reading test discrimination was very high in almost all of the testlets in both the pre- and post-tests, and the reliability of the scores was good in the pre-test and excellent in the post-test. In the writing test, most of the testlets showed small or very small discrimination indices in the pre-test, and the reliability of the scores was adequate with some deficiencies. Furthermore, the discrimination indices were average and reliability was excellent in the post-test.

On the other hand, great improvements were found in the post-test in both reading and writing. Regarding age, there were differences in reading and writing between the 4 and 5 year olds in the pre-test and in the post-test. In addition, improvements in the post-test were higher for the 5 year olds in both tests. However, with respect to sex, no differences were found in reading and writing between the boys and girls in the pre-test or the post-test, as the improvements observed were similar for both.

Discussion y Conclusions

This study analyzed the performance and improvement in reading and writing of preschool students, according to age and sex, after analyzing the psychometric properties of the tests used.

The differences found in reading and writing among the pre-test and post-test scores of the 4- and 5-year-old children, and between both age groups, indicate that

although the teaching of literacy skills is not compulsory in many countries at these early ages, the children do benefit when taught these skills early on (Brand & Dalton, 2012, González & Delgado, 2006, 2009, González et al., 2011, 2012, Lovett, et al., 2017, Slavin & Madden, 2001, Suggate, et al. , 2013; Park, et al., 2015; Wright et al., 2008).

The teaching of reading and writing in some countries takes place before the children start primary school (Slavin & Madden, 2001, Slavin, et al., 1996). However, in Spain, in particular, according to current legislation, the teaching of reading and writing is introduced from the onset of primary education, although in preschool an approach to written language is established through texts related to everyday life, where language is used as an instrument of communication and representation (three-five years). (CECJA, 2008, MEC, 2006). Some of the contents that are given at these ages are related to differentiating between written forms and other forms of graphic expression, the conventions of the writing system (linearity, arbitrariness, orientation, and organization), the identification of letters and words and important phrases such as the child's own name and the names of the people in the child's environment, the understanding of what was read and the expression of short messages (shopping list or congratulations). Although the legislation does not require systematic and in-depth learning of reading and writing until the beginning of primary education, teachers tend to introduce these concepts to their students in preschool due to the benefit it has on their academic performance. Although this study does not control the amount and type of instruction that the students received, the results do show, however, an improvement in reading and writing between the children in the second and third year of preschool. In this regard, our findings are in line with those positions that opt for early literacy, since our results show benefits in teaching literacy skills at these ages. (Brand & Dalton, 2012, Brown et al., 2012, Dunphy, 2012, Elliott & Olliff, 2009, González et al., 2011, 2012, González-Valenzuela & Martín-Ruiz, 2017, Lovett, et al., 2017 Slavin & Madden, 2001; Suggate, et al., 2013; Park, et al., 2015; Vadasy & Sanders, 2008; Wright et al., 2008). Most of these studies highlight that some of the benefits of the early intervention of the teaching and learning of reading and writing lie in obtaining a better academic performance in the short and long term and in a decrease in the number of cases of children with learning difficulties.

Regarding the differences in performance in reading and writing associated with sex, our results differ from some studies carried out in primary and secondary education (Camarata & Woodcock, 2006, Gomendio, 2015, Logan & Johnston, 2009, Martínez García & Córdoba, 2012, Mullis et al., 2003, 2007, MECD, 2016, 2017, OECD 2010a, b, Pargulski & Reynolds, 2017, Reynolds, et al., 2015, Schwabe, et al., 2014, Troia et al., 2013). It could be thought that if there are no sex-related differences in reading and writing at early ages then perhaps the gap between boys and girls at later ages could be due to the influence of sociocultural and educational contexts on the neurological development of children (Wood & Eagly, 2002). Another possible explanation lies in the influence of motivation and attitude for reading and writing on the performance of written language (Logan & Johnston, 2009, Martínez García & Córdoba, 2012, Pargulski & Reynolds, 2017, Schwabe, et al., 2014; Troia et al., 2013). Therefore, it is important for schools to promote early intervention programs that would optimize the teaching-learning processes of reading and writing in order to help eliminate the differences found between the two sex groups, fostering interest in written language from early ages.

Finally, this study shows the importance of having quality instruments that assess the learning of reading and writing at an early age. This is in line with what other

authors have highlighted regarding the need for proving the psychometric quality of the language assessment measures currently available (Denman, Speyer, Munro, Pearce, Chen, & Cordier, 2017). However, the lower number of children participating in the post-test as opposed to the pre-test due to changes in the school should be taken into account, as well as the fact that we were unable control the amount and type of instruction that the students received in each school or the sociocultural influences on performance in reading and writing according to age and sex.

Future research is required to go deeper into this area, improving the tests aimed at assessment in early childhood, creating scales in different cultures and languages of different orthographic consistency. The analysis of the testlets representing the components of reading and writing, such as accuracy and reading comprehension and accuracy in copying and dictation, should be addressed in future studies. Additionally, the analyses should be carried out during first and second grade in order to determine the extent to which the results observed in preschool persist.

References

- Brand, S.T., & Dalton, E.M. (2012). Universal Design for Learning: Cognitive Theory into Practice for Facilitating Comprehension in Early Literacy. *Forum on Public Policy Online*, 1, 1-19.
- Brown, R., Scull, J., Nolan, A., Raban, B., & Deans, J. (2012). Young Learners: Mapping the Beliefs and Practices of Preschool Teachers in Relation to Early Literacy Development. *Australian Educational Research*, 39(3), 313-331. <http://dx.doi.org/10.1007/s13384-012-0061-0>
- Camarata, S. & Woodcock, R. (2006). Sex Differences in Processing Speed: Developmental Effects in Males and Females. *Intelligence*, 34(3), 231-252. <http://dx.doi.org/10.1016/j.intell.2005.12.001>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New Jersey: Lawrence Erlbaum Associates.
- Consejería de Educación de la Junta de Andalucía (2008). Orden de 5 de Agosto, por la que se desarrolla el Currículo correspondiente a la Educación Infantil en Andalucía. *Boletín oficial de la Junta de Andalucía*, No. 169. Sevilla.
- Denman, D., Speyer, R., Munro, N., Pearce, W., Chen, Y., & Cordier, R. (2017). Psychometric Properties of Language Assessments for Children Aged 4–12 Years: A Systematic Review. *Frontiers of Psychology*, 8, 1515. <http://dx.doi.org/10.3389/fpsyg.2017.01515>
- Dunphy, E. (2012) Children's participation rights in early childhood education and care: The case of early literacy learning and pedagogy. *International Journal of Early Years Education*, 20(3), 290-299. <http://dx.doi.org/10.1080/09669760.2012.716700>
- Elliott, E.M., & Olliff, C.B. (2009). Developmentally Appropriate Emergent Literacy Activities for Young Children: Adapting the Early Literacy and Learning Model. *Early Childhood Educational Journal*, 35, 551–556. <http://dx.doi.org/10.1007/s10643-007-0232-1>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191. <http://dx.doi.org/10.3758/BF03193146>
- Gomendio, M. (2015). Igualdad de género en la evaluación PISA. Secretaria de Estado de Educación, Formación Profesional y Universidades.

- González, M.J., & Delgado, M. (2006). Enseñanza-aprendizaje de la lectura y la escritura en educación infantil 4 años. *Revista de Psicología General y Aplicada*, 59(4), 465-478.
- González, M.J., & Delgado, M. (2009). Rendimiento académico y enseñanza-aprendizaje de la lectoescritura en Educación Infantil y Primaria: un estudio longitudinal. *Infancia y Aprendizaje*, 32(3), 265-276. <http://dx.doi.org/10.1174/021037009788964114>
- González-Valenzuela, M.J., & Martín-Ruiz, I. (2017). Effects on Reading of an Early Intervention Program for Spanish Children at Risk of Learning Difficulties. *Remedial and Special Education*, 38(2), 67-75. <http://dx.doi.org/10.1177/0741932516657652>
- González, M. J., Martín, I., & Delgado, M. (2011). Intervención temprana de la lectoescritura en sujetos con Dificultades de Aprendizaje. *Revista latinoamericana de Psicología*, 43(1), 35 – 44.
- González, M.J., Martín, I., & Delgado, M. (2012). Teaching literacy and decreased risk of learning disabilities. *Revista de Psicodidáctica*, 17(2), 253-270. <http://dx.doi.org/10.1387/RevPsicodidact.4502>
- Hernández, A., Ponsoda, V., Muñiz, J., Prieto, G., & Elosua, P. (2016). Revisión del modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 37, 161-168.
- Logan, S. & Johnston, R. (2009). Gender Differences in Reading Ability and Attitudes: Examining where These Differences Lie. *Journal of Research in Reading*, 32(2), 199-214. <http://dx.doi.org/10.1111/j.1467-9817.2008.01389.x>
- Lovett, M.W., Frijters, J.C., Wolf, M., Steinbach, K.A., Sevcik, R.A., & Morris, R.D. (2017). Early intervention for children at risk for reading disabilities: The impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology*, 109(7), 889-914. doi:10.1037/edu0000181
- Martínez García, J.S., & Córdoba, C. (2012). Rendimiento en lectura y género: una pequeña diferencia motivada por factores sociales (pp. 144-179). En Instituto Nacional de Evaluación Educativa (Ed). PIRLS - TIMSS 2011. Estudio Internacional de progreso en comprensión lectora, matemáticas y ciencias. IEA Volumen II: Informe Español. Análisis Secundario. Madrid: MEC.
- Ministerio de Educación de España (2006). Decreto 1630/2006, de 29 de diciembre por el que se establecen las enseñanzas mínimas del segundo ciclo de Educación Infantil. *Boletín oficial de España*, No. 4. Madrid.
- Ministerio de Educación, Cultura, & Deporte (2017). [PIRLS 2016. Estudio internacional de progreso en comprensión lectora. IEA. Informe español](#). Secretaria de Estado de Educación, Formación Profesional y Universidades. Madrid.
- Ministerio de Educación, Cultura, & Deporte (2016). Pisa 2016. Programa para la evaluación internacional de los alumnos. Informe español. Secretaria de Estado de Educación, Formación Profesional y Universidades. Madrid.
- Mullis, I., Martin, M., Kennedy, A., & Foy, P. (2007). PIRLS 2006 International Report. IEA's Progress in International Reading Literacy Study in Primary Schools in 40 countries. TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.
- González, E., & Kennedy, A. (2003). PIRLS 2001 International Report. IEA's Study of Reading Literacy Achievement in Primary Schools in 35 countries. International Association for the Evaluation of Educational Achievement - International Study Center, Lynch.
- OCDE (2010a). PISA 2009. Learning to learn. Vol. III. Paris: OECD Publishing.

- OCDE (2010b). PISA 2009. What Students Know and Can Do. Vol. I. Paris: OECD Publishing.
- Pargulski, J. R., & Reynolds, M.R. (2017). Sex differences in achievement: Distributions matter. *Personality and Individual Differences* 104 (2017) 272–278. Doi: 10.1016/j.paid.2016.08.016
- Park, Y., Chaparro, E. A., Preciado, J., & Cummings, K. D. (2015). Is earlier better? mastery of reading fluency in early schooling. *Early Education and Development*, 26(8), 1187-1209. doi:10.1080/10409289.2015.1015855
- Raju, N.S. (1977). A generalization of coefficient alpha. *Psychometrika*, 42(4), 549-565
- Reynolds, M.R., Scheiber, C., Hajovsky, D.B., Schwartz, B., & Kaufman, A.S. (2015). Gender Differences in Academic Achievement: Is Writing an Exception to the Gender Similarities Hypothesis? *The Journal of Genetic Psychology*, 176(4), 211–234, 2015.
- Revuelta, F., & Guillén, A. (1987). Madurez y edad inicial para la lectura. Comunicación presentada a la 5ª CEL, 154-155.
- Schwabe, F., McElvany, N., & Trendtel, M. (2014). The School Age Gender Gap in Reading Achievement: Examining the Influences of Item Format and Intrinsic Reading Motivation. *Reading Research Quarterly*, 50(2), 219–232. doi: 10.1002/rrq.92
- Slavin, R.E., & Madden, N.A. (2001). One million children: Success for all. California: Corwin.
- Slavin, R.E., Madden, N.A., Dolan, L.J., & Wasik, B.A. (1996). Every child, every school: Success for all. Thousand Oaks, CA: Corwin Press.
- Speer, D.C. (1992). Clinical significant change: Jacobson and Traux (1991) revisited. *Journal of Consulting & Clinical Psychology*, 60(3), 402-408.
- SPSS, Inc. (2006). SPSS 14.0 Base User's Guide. Chicago, IL: Author.
- Suggate, S.P., Schaughency, E. A., & Reese, E. (2013). Children learning to read later catch up to children reading earlier. *Early Childhood Research Quarterly*, 28(1), 33-48. doi:10.1016/j.ecresq.2012.04.004
- Troia, G.A., Harbaugh, A.G., Shankland, R.K., Wolbers, K.A., & Lawrence, A.M. (2013). Relationships between Writing Motivation, Writing Activity, and Writing Performance: Effects of Grade, Sex, and Ability. *Reading and Writing: An Interdisciplinary Journal*. 26,(1), 17-44. <http://dx.doi.org/10.1007/s11145-012-9379-2>
- Vadasy, P.F., & Sanders, E.A. (2008). Repeated reading intervention: outcomes and interactions with reader's skills and classroom instruction. *Journal of Educational Psychology*, 100(2), 272-290. <http://dx.doi.org/10.1037/0022-0663.100.2.272>
- Wood, W., & Eagly, A.H. (2002). A cross-cultural analysis of the behavior of women and men: implications for the origins of sex differences. *Psychological Bulletin*, 128(5), 699-727. <http://dx.doi.org/10.1037/0033-2909.128.5.699>
- Wright J. A., Stackhouse J., & Wood J. (2008). Promoting language and literacy skills in the early years: lessons from interdisciplinary teaching and learning. *Child Language Teaching and Therapy*, 24(2), 155–171. <http://dx.doi.org/10.1177/0265659007090292>