

DYSLEXIA

Published online 30 September 2013 in Wiley Online Library  
(wileyonlinelibrary.com). DOI: 10.1002/dys.1462

# ■ A new procedure to measure children's reading speed and accuracy in Italian

Isabella Morlini<sup>1\*</sup>, Giacomo Stella<sup>2</sup> and Maristella Scorza<sup>2</sup><sup>1</sup>*Department of Economics, University of Modena and Reggio Emilia, Via Berengario 51, 41121 Modena, Italy*<sup>2</sup>*Department of Social Sciences, University of Modena and Reggio Emilia, 42121 Reggio Emilia, Italy*

Impaired readers in primary school should be early recognized, in order to assess a targeted intervention within the school and to start a teaching that respects the difficulties in learning to read, to write and to perform calculations. Screening procedures, inside the primary schools aimed at detecting children with difficulties in reading, are of fundamental importance for guaranteeing an early identification of dyslexic children and reducing both the primary negative effects—on learning—and the secondary negative effects—on the development of the personality—of this disturbance. In this study, we propose a new screening procedure measuring reading speed and accuracy. This procedure is very fast (it is exactly 1 min long), simple, cheap and can be provided by teachers without technical knowledge. On the contrary, most of the currently used diagnostic tests are about 10 min long and must be provided by experts. These two major flaws prevent the widespread use of these tests. On the basis of the results obtained in a survey on about 1500 students attending primary school in Italy, we investigate the relationships between variables used in the screening procedure and variables measuring speed and accuracy in the currently used diagnostic tests in Italy. Then, we analyse the validity of the screening procedure from a statistical point of view, and with an explorative factor analysis, we show that reading speed and accuracy seem to be two separate symptoms of the dyslexia phenomenon. Copyright © 2013 John Wiley & Sons, Ltd.

**Keywords:** internal consistency; learning disorder; primary school; screening procedure

## INTRODUCTION

Nowadays, a prevention approach to reading problems is well-established: schools do not wait for a student to fail, but they screen all students to identify those who, despite a normal education programme, show difficulties from the beginning of reading acquisition. To have any chance of escaping this adverse path, students must obtain help as soon as possible, and for this reason, it is very important to have instruments capable of identifying children who are at risk at a very early stage. Studies have shown that reading problems become increasingly more resistant to intervention and treatment after third grade (Snow, Burns & Griffin, 1998). Designing effective early reading screening instruments is complicated as the time of year and age/grade at which they are administered may greatly

\*Correspondence to: Isabella Morlini, Department of Economics, University of Modena and Reggio Emilia, Via Berengario 51, 41121 Modena, Italy. E-mail: [isabella.morlini@unimore.it](mailto:isabella.morlini@unimore.it)

influence the predictive value of a test, even within a short time frame or the same school year (O'Connor & Jenkins, 1999). The appropriate time to begin screening young children has been debated for a long time. In some cases, the literature argues that for young children, screening of early literacy skills should occur before formal literacy instruction begins (Badian, 1982; 2000; Invernizzi, Justice, Landrum & Booker, 2004; Justice, Invernizzi & Meier, 2002). In regular orthographic systems, such as in Italy, Spain and Finland, at the end of first grade, children are typically close to the ceiling in terms of reading accuracy (Goswami, Gombert & Fraca de Barrera, 1998; Seymour, Aro & Erskine, 2003). Italian studies on reading acquisition show that at the beginning of primary school, children are highly heterogeneous, which is caused by different levels of maturity and different cognitive and linguistic abilities. In kindergarten, Italian children are not trained systematically in reading acquisition and phonology, and this leads to a dishomogeneous situation between children at the beginning of first grade (Allamandri, 2005; Riccardi Ripamonti, Cividati, Truzoli & Russo, 2007; Stella & Apolito, 2004). Nevertheless, because of the consistency of Italian orthography, at the end of first grade, the large majority of children successfully perform in reading and spelling (Stella, 2004).

In Italy, most primary school screenings typically use group tests, whereby all students are examined at the same time (Bellocchi, 2011; Cornoldi & Colpo 1986; 1998; Faglioni, Gatti, Paganoni & Robutti, 1967; Stella & Apolito 2004). Even if a lot of individual tests are available in Italian (Cornoldi & Colpo, 1998), teachers seem to prefer group tasks, as this procedure is considered more economic and time saving. However, also group tests face some criticism, as children may score poorly because of behaviour, attention and task motivation problems (Rathvon, 2004; Scarborough, 1998). Moreover, group tasks do not allow each individual to read aloud so that reading has to be tested with different print skills performed silently (letter and word or pseudo-word matching, word comprehension matching words and pictures in multiple choice format). Reading is a multidimensional skill, and screens should reflect this, through reading aloud, word reading and reading fluency. When studying reading acquisition, it is useful to analyse decoding skills and comprehension processes separately, as two different levels of processes are involved (Ransby & Swanson, 2003).

Our aim was to use a single task to examine reading skill in the most similar way to the natural decoding process, that is, reading aloud. To this purpose, we decided to use a 1-min reading test to balance economic issues, in terms of time administration cost, with the sensitivity and specificity of the task. Some researchers claim that screening battery (multiple measures) obtain better classification accuracy (Jenkins & O'Connor, 2002), but the problem of administration time cost per person still remains. If a single task is used, you can get advantages in administration time cost per person, but you need to carefully consider high concurrent validity between the screen task and a benchmark battery of standard test to assess reading skills. Screening should occur every year across the primary grades. One-minute reading procedure was designed by Nicolson and Fawcett (1997) and is frequently used in research both to measure reading performance specifically in children and adults (Nicolson & Fawcett, 1997; Willburger & Landerl, 2009) and to match reading performance with some other skill such as perceptual attention (Schulte-Krue, 2001).

Drawing from the results obtained by administering our screening procedure and, as a benchmark, two currently used reading tests, to about 1500 students in primary school in Italy, the purposes of this work are as follows:

- To study the empirical distributions of the variables used in the suggested screening procedure and in the benchmark tests and to analyse their discriminative power in detecting impaired readers;
- To study the bivariate relationships between variables, in order to investigate if the reading speed and the reading accuracy measured in a 1-min procedure are correlated with the reading speed and the reading accuracy, respectively, measured by means of longer tests; and
- To study the validity of the screening procedure, through an explorative factor analysis and through the estimate of the internal consistency.

The paper is organized as follows. In the New Screening Procedure and the Benchmark Tests section, we briefly illustrate the new screening procedure and the tests used as benchmarks. In the Method and Subjects section, we report univariate and bivariate statistics of the variables involved in the study, we analyse the empirical distributions of these variables and we discuss the choice of the threshold values for classifying the student's performance. In the Univariate and Bivariate Analyses section, we investigate the validity of the new screening procedure. Finally, in the Internal Consistency section, we give some concluding remarks.

## THE NEW SCREENING PROCEDURE AND THE BENCHMARK TESTS

The new screening procedure is called SPILLO ('tool for identifying slowness in reading aloud'). It is implemented within a software, and thus, the results of the screening are immediately available after administration. It is designed for Italian students in primary school, from age 6 years (first grade) to 10 years (fifth grade). Usually, reading tests are 'paper and pencil'; that is, while a student reads, the examiner times the reading and marks the mistakes. At the end of the test, the examiner counts the errors and calculates the time and number of syllables read in a second. Eventually, he/she also computes the z-scores of errors and speed (syllables per second) in order to classify single performance. Then, the examiner classifies the student as an 'impaired reader' or as a 'not impaired reader' on the basis of the normative threshold values for each indicator. This procedure is very simple for experts, but it may result rather difficult for examiners who lack a statistical background. Indeed, results are usually affected by many errors due to erroneous calculus or wrong interpretation of the results. In SPILLO scoring and recording, mistakes are avoided by the use of the computer, but ecological situation to read a paper is maintained for the subjects. The student is asked to read a text for exactly 1 min. The examiner clicks the word spelled wrongly on the screen and, at the tone produced after 1 min, he stops the reader. The examiner has just to click the last word read by the student to obtain automatically the number of words, the number of syllables read in a second, the number of errors and, if the case, the z-scores. For our purposes, a text reading test was chosen because in a daily reading assignment the reader faces a text and not a list of isolated words. Moreover, when reading a text, semantic and morphosyntactic information can enhance the reader's

understanding, whereas reading a list of words or pseudo-words is primarily a measure of intact or impaired lexical or sub-lexical processes. Finally, reading a text allows the evaluation of the efficiency of the so-called instrumental reading, that is, the ability to quickly and correctly recognize and name words within a text (De Beni, Cisotto & Carretti, 2001). The chosen text tells a story. This kind of task allows testing of the processes involved in reading meaningful material in a full and complete manner. For this study, a story was selected from the pages of a primary school text book (Lastrego & Testa, 1995). The list of words and nonwords (Sartori, Job & Tressoldi, 1995) was used to establish concurrent validity between SPILLO and standardized measures of reading skills. MT reading text (Cornoldi & Colpo, 1998) was excluded as it is a similar task to SPILLO, and word and nonword decoding test was preferred as it is considered the best benchmark of reading impairment (Hermann, Matyas and Pratt, 2006).

## METHOD AND SUBJECTS

### The Text Reading Task

The criteria used to choose the text was reading difficulty as scored by the Gulpease index. The Gulpease index is an evolution of the Flesch index (Flesch, 1948) and was adapted to the Italian language by Franchina and Vacca (1986). This index compares the words in the text with those in the base vocabulary of the Italian language (VdB) of De Mauro (1997) ([www.eulogos.net/it/censor](http://www.eulogos.net/it/censor)). The Gulpease index uses a predefined scale from 0 (minimum readability) to 100 (maximum readability). This scale is based on the evaluation of real understanding of a body of texts by readers of three different school ages. According to this method, the text chosen 'Ho paura' [I am afraid] has a score of 71. Therefore, it belongs to the group of materials considered 'difficult to read' for the school group tested (first grade students). However, out of the total of 181 words, 173 belong to the VdB (95.57%) and only eight words are not in the VdB (4.41%). When analysing some primary school textbooks, the average readability score of the selected texts was found to be below 60 (Piemontese & Cavaliere, 1997). A score of 60 is indicated as the limit above which middle school students have no difficulties in understanding the content of the text, when they read it. Therefore, these scores indicate that the child is not capable of reading and understanding the text in his textbook by himself. Because the text is a new test and its value must be empirically determined, all the subjects of our sample were also evaluated using two more reading tests, which are used to assess dyslexia. These two tests consist of a list of single words and a list of pseudo-words to read aloud, standardized for Italian students (DDE-2, Sartori, Job & Tressoldi, 1995; 2007).

In SPILLO, the variables monitored are as follows:

- $Y_1$ : number of words read in a minute;
- $Y_2$ : number of syllables per second read in a minute; and
- $Y_3$ : number of wrong spelling in a minute.

The choice to use the same text for all grades was made for different reasons. First, a unique test for all primary grades is more practical for the administrator (namely teachers) and reduces methodological errors. Second, a developmental trend could be obtained and compared with the benchmark of word and pseudo-word reading.

Third, as there is much data showing that in transparent orthographies, semantic, orthographic and perceptual features of the written text (such as font size.) are irrelevant to differentiate reading performance along different school grades. Tressoldi (2008) demonstrated that Italian readers improve only in speed along ages, as differences in readability on speed and accuracy due to lexical and orthographic features do not vary across age and school grade. Zoccolotti, De Luca, Di Filippo, Judica and Martelli (2008), when looking for crucial factors in development of reading speed, demonstrated that length, word frequency and lexicality did not account for a major portion of variance, whereas a global factor is more relevant in developing reading speed. We decided not to consider sensitivity to stress assignment, drawing from the study of Holliman, Wood and Sheehy (2010), which shows that this sensitivity is not essential to the realization of reading in syllable-timed languages, whereas it seems to be more related to phonological awareness and reading abilities in stress-timed languages. The association between stress and word reading is relevant as the ability to manipulate stressed and unstressed syllables seems to help clarification of ambiguous phonemes and to facilitate word recognition in reading (Wood & Terrell, 1998). In Italian, reading ambiguity is limited to stress assignment on words that are identical in phonological and graphemic structure but differ on stress position. Stress is graphically represented only if it occurs on the last syllable, whereas within the word, it is never represented. Italian readers have to assign stress bearing on semantic aspects. This is a very rare occasion (for example, the words ‘ancora’ [anchor] and ‘anc’ora’ [more]); therefore, it was not difficult to avoid this occurrence in our text (Angelelli et al., 2004).

### The Benchmark Test

The benchmark tests are two currently used ‘paper and pencil’ diagnostic tests in Italy. While the student reads, the examiner times the reading and makes a note of the mistakes. During the first test, the student is asked to read a list of words and during the second test a list of pseudo-words. These two lists have been introduced and studied by Sartori, Job and Tressoldi (1995; 2007). In the two benchmark tests, the variables monitored are as follows:

- $X_1$ : time (in seconds) in reading the list of words;
- $X_2$ : number of syllables per second read in the list of words;
- $X_3$ : number of wrong spellings in reading the list of words;
- $X_4$ : time (in seconds) in reading the list of nonwords;
- $X_5$ : number of syllables per second read in the list of nonwords; and
- $X_6$ : number of wrong spellings in reading the list of nonwords.

The new screening procedure and the two benchmark tests have been administered to all subjects of our sample.

### Subjects

One thousand four hundred sixty-nine children are attending primary school in the Lombardia and the Emilia Romagna regions (Northern Italy). The test was administered during the month of February to students in grades 2 (mean age

7.5 years), 3 (mean age 8.3 years), 4 (mean age 9.3 years) and 5 (mean age 10.1 years). First graders (mean age 6.8 years) were tested during the month of May in order to achieve a more uniform reading level at the end of the first year of formal teaching. On the other hand, it is known that because of the regularity of Italian orthography, by the end of the first year of primary school, 90% of the children are able to read (Cossu, 1999; Orsolini, Fanari & Maronato, 2005; Stella, 2004; Stella & Apolito, 2004). Foreign children and children with learning disabilities (according to Law n. 104/92) were not included in our sample. Parents of participants or legal guardians gave informed consent.

The composition of the sample is 333 students in first grade, 384 students in second grade, 200 students in third grade, 276 students fourth grade and 276 students in fifth grade. We selected 75 schools for the study in the Emilia Romagna and Lombardia regions (Northern Italy), and we administered the tests and the screening procedure to all students attending these schools. Therefore, the number of participants in each grade was decided a priori but because of methodological reasons, some subjects were excluded from the sample. The fact that more students are in first and second grades is convenient for our study, because more variability in reading performances is present in first grades of primary school.

#### **Procedure to Administer the SPILLO Test**

Test has to be administered in a quiet environment, outside the classroom. The text to read is printed on a cardboard in black and white at high contrast, in Times New Roman font and is spaced 1.5. Each student is asked to read the text aloud as quickly as possible but as accurate as possible at the same time. The examiner puts the cardboard in front of the student and read the title of the text pointing to where the student has to start reading. The student is asked to read until a stop is given by the examiner. While the subject is reading on his cardboard, the examiner follows the reading on a computer screen, clicking on a start label to measure reading time and clicking on the words spelled wrongly. When the experimenter clicks on a word to enter a mistake, a window pops up with three options: wrong, auto-correction and omitted, and the correct option can be chosen in order to classify errors for a further qualitative evaluation. This procedure does not interfere with the subject reading, as the computer screening is not on the student visual field. After 1 min from the start, a tone is produced by the computer and the examiner interrupts the reading. Then examiner clicks the last word read by the student and obtains automatically, on a window, the number of words, the number of syllables read per second, the number of errors and, if the case, the z-scores.

## **UNIVARIATE AND BIVARIATE ANALYSES**

### **Reading Speed**

In Table I, we list the values of some univariate statistics of the variables measuring the reading speed ( $X_1$ ,  $X_2$ ,  $X_4$ ,  $X_5$ ,  $Y_1$ , and  $Y_2$ ). Figure 1 shows the empirical distribution of these variables through boxplots. Performances in the reading speed improve from first grade to fifth grade: both the median and the mean values of  $X_1$  and  $X_4$  decrease whereas the mean and the median values of  $X_2$ ,  $X_5$ ,  $Y_1$  and  $Y_2$  increase.

Table 1. Univariate statistics obtained in the sample for variables measuring the reading speed

	X <sub>1</sub>					X <sub>4</sub>					Y <sub>1</sub>				
	I	II	III	IV	V	I	II	III	IV	V	I	II	III	IV	V
N	333	384	200	276	276	333	384	200	276	276	333	384	200	276	276
Min	92	69	52	50	43	58	49	33	34	31	8	21	48	9	57
Max	1148	595	289	348	206	542	311	210	234	118	127	157	164	180	180
X <sub>0.05</sub>	145	96	70	59	54	89	68	52	49	38	16	34	64	78	96
Median	314	177	110	89	74	151	110	84	72	59	41.0	66.0	97.0	125.0	146.0
X <sub>0.95</sub>	723	381	209	153	121	342	206	157	109	97	72	113	138	170	180
Mean	360	198	121	96	79	173	120	88	74	63	43.0	69.0	97.7	123.5	143.7
Standard deviation	190.5	88.9	43.7	34.2	23.5	82.4	45.4	29.8	23.8	17.8	19.9	23.9	24.3	26.9	25.3
Coefficient of variation	0.53	0.45	0.36	0.36	0.30	0.48	0.38	0.34	0.32	0.28	0.46	0.35	0.25	0.22	0.18
Skewness	1.66	1.49	1.30	2.53	1.84	1.76	1.64	1.45	2.23	0.61	1.21	0.64	-0.23	-0.42	-0.76
Kurtosis	3.29	2.84	1.82	12.30	5.47	3.81	3.46	2.69	11.51	0.02	2.94	0.44	-0.65	0.86	0.36
								X <sub>5</sub>					Y <sub>2</sub>		
Min	0.24	0.47	0.97	0.80	1.36	0.23	0.41	0.60	0.54	1.07	0.30	0.78	1.77	0.32	2.07
Max	3.04	4.06	5.38	5.60	6.51	2.17	2.57	3.82	3.71	4.06	4.23	5.25	5.53	5.98	5.98
X <sub>0.05</sub>	0.39	0.73	1.34	1.83	2.31	0.37	0.61	0.80	1.16	1.30	0.58	1.27	2.30	2.80	3.32
Median	0.89	1.58	2.56	3.16	3.78	0.83	1.15	1.51	1.76	2.14	1.48	2.38	3.33	4.18	4.87
X <sub>0.95</sub>	1.93	2.92	4.00	4.75	5.19	1.42	1.85	2.43	2.82	3.32	2.65	3.82	4.55	5.70	5.98
Mean	0.98	1.67	2.59	3.20	3.77	0.86	1.18	1.59	1.85	2.17	1.56	2.46	3.36	4.16	4.82
Standard deviation	0.48	0.66	0.81	0.89	0.92	0.33	0.37	0.49	0.53	0.62	0.68	0.77	0.73	0.81	0.79
CV	0.49	0.40	0.31	0.28	0.24	0.38	0.32	0.31	0.29	0.29	0.44	0.31	0.22	0.19	0.16
Skewness	1.48	0.74	0.34	0.16	0.03	0.54	0.45	0.83	0.72	0.69	1.06	0.45	-0.09	-0.34	-0.69
Kurtosis	3.32	0.41	-0.16	-0.22	-0.18	0.38	0.24	1.87	0.65	0.17	2.10	0.20	0.68	1.04	0.05

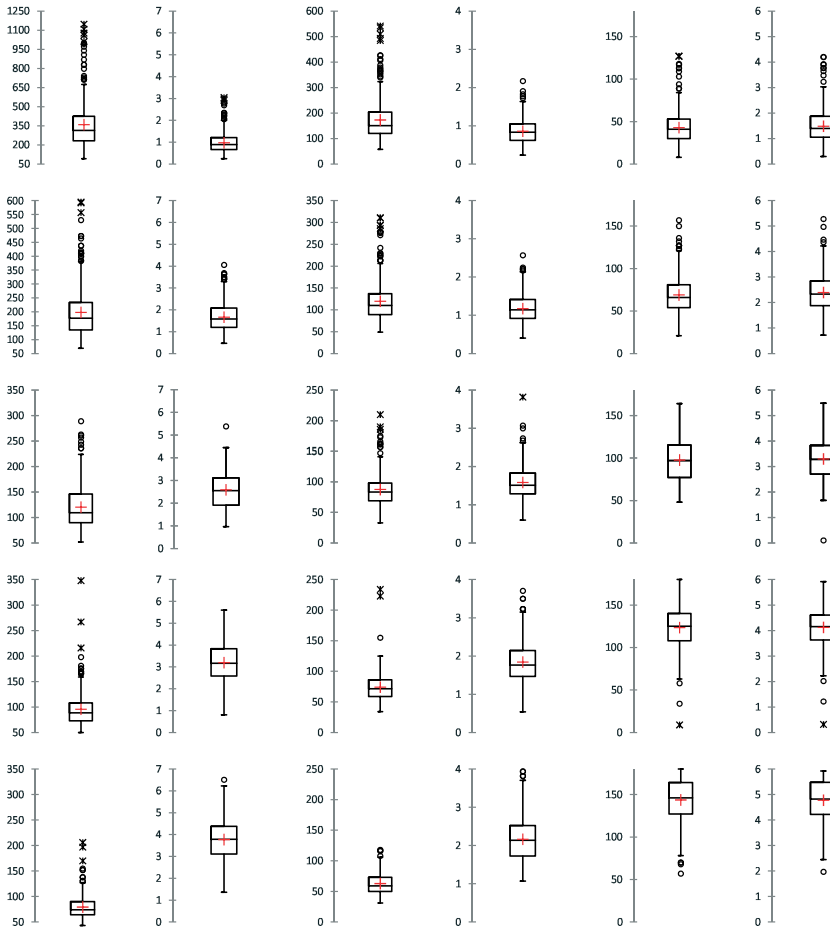


Figure 1. Boxplots of the following variables (from the left):  $X_1$ ,  $X_2$ ,  $X_4$ ,  $X_5$ ,  $Y_1$  and  $Y_2$ . The first line reports boxplots in class I, the second line in class II, the third line in class III, the fourth line in class IV and the fifth line in class V.

Variables measuring the number of words and the number of syllables read in a second have a similar pattern: the average values of  $Y_1$  and  $Y_2$  across the five grades have a behaviour similar to the average values of  $X_2$  and  $X_5$ . Dispersion, measured by the coefficient of variation, always decreases with the grade level. In the time of reading ( $X_1$  and  $X_4$ ) and in  $Y_1$  and  $Y_2$ , a drop of the coefficient of variation corresponding to third grade is evident (Figure 2). This drop is not present in  $X_2$  and  $X_5$ . The larger variability in first, second and third grades may be explained by considering that many covariates (such as the cultural level and the experiences in day nursery) have a great influence on the reading performances. From third grade, in general, these covariates become less important, and the scholastic population is more homogeneous. Variables measuring the reading speed in the benchmark tests ( $X_1$ ,  $X_2$ ,  $X_4$ , and  $X_5$ ) have a positive skew and present outlying values higher than  $x_{0.75} - 1.5(x_{0.75} - x_{0.25})$ , in all grades. These characteristics are desirable for  $X_1$  and  $X_4$ , which have a positive direction of pathology (impaired readers are children with high values in these variables), but not for  $X_2$  and  $X_5$ , which have a negative direction of pathology (impaired readers are children having small values in these variables).



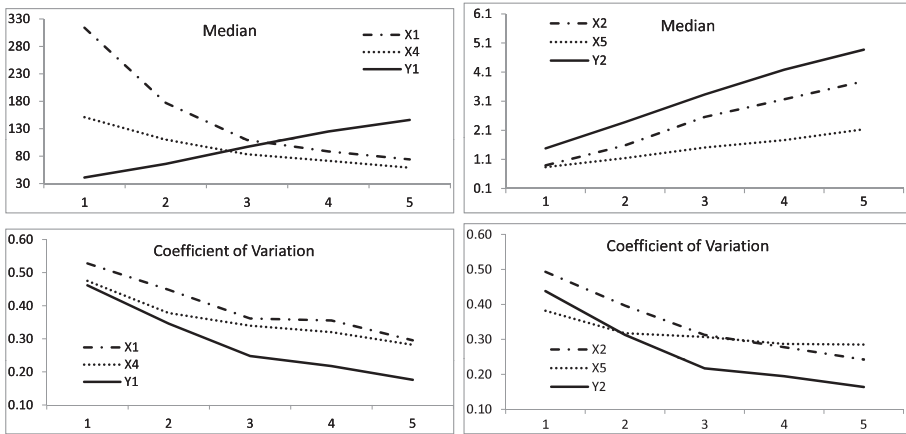


Figure 2. Coefficient of variation and median values from class I to class V.

For children having outlying values in  $X_1$  and  $X_4$ , more accurate analyses are needed to investigate if this impairment is due to dyslexia or to other learning disabilities. In general, the aim of the screening is to detect impaired readers, and subsequent psychological and medical tests are needed to investigate the cause of this impairment.  $Y_1$  and  $Y_2$  have a positive skew in first and second grades but a negative skew in third, fourth and fifth grades. In fourth and fifth grades, these variables have also outlying values smaller than  $x_{0.25} + 1.5(x_{0.75} - x_{0.25})$ . These features show that  $Y_1$  and  $Y_2$  have a more discriminative power during the last grades of primary school, when the students read more fluently. The kurtosis indicates that all variables, in each grade, are far from normality. With the nonparametric tests of Shapiro-Wilk, Anderson-Darling and Jarque-Bera, the null hypothesis of Gaussian distribution is rejected, even for  $\alpha = 0.001$ .

The currently used thresholds for  $X_1$ ,  $X_2$ ,  $X_4$ , and  $X_5$  are based on the assumption of normality. The thresholds are used to classify students as normal readers or impaired readers. They have been specified on the basis of the mean and the variance, assuming a normal distribution (Sartori, Job & Tressoldi, 1995; 2007). The thresholds have been obtained as  $\mu + 2\sigma$  (for  $X_1$  and  $X_4$ ) and as  $\mu - 2\sigma$  (for  $X_2$  and  $X_5$ ), where  $\mu$  indicates the mean and  $\sigma$  the standard deviation, considering that in a Gaussian distribution these values both exclude about 2% of the population. The estimated values of  $m$  and  $s$ , reported in Sartori et al. (2007) and currently used as normative values in Italy, have been estimated only for second, third, fourth and fifth grades. Using the  $t$ -test for the means and the nonparametric test of Levene for the variances, they result significantly different ( $\alpha = 0.05$ ) from the means and the variances obtained in our study and reported in Table I. The causes of these differences are, at least, two-fold: (i) the size of the samples is different (our sample is much larger); and (ii) outliers have been removed in the normative data before computing the mean and the standard deviation. We do not think it is a good practice to remove outliers, when these values are not erroneous. Because of the skewness and the presence of outliers, the median and the percentiles should be used instead of the mean and the variance, in order to describe the variables and to define the thresholds. Indeed, the percentiles are more robust to outliers. The normative thresholds, in our sample, lead to the following percentages of students classified as impaired readers:

Table 2. Correlation matrices for variables measuring the reading speed

First grade							Fourth grade						
	X1	X2	X4	X5	Y1	Y2		X1	X2	X4	X5	Y1	Y2
X1	1	-0.79	0.91	-0.81	-0.72	-0.73		1	-0.89	0.84	-0.71	-0.76	-0.75
X2	-0.79	1	-0.72	0.9	0.91	0.91	X1	1	1	-0.77	0.83	0.77	0.76
X4	0.91	-0.72	1	-0.86	-0.66	-0.66	X2	-0.89	1	1	-0.88	-0.61	-0.61
X5	-0.81	0.9	-0.86	1	0.82	0.82	X4	0.84	-0.77	1	1	0.61	0.6
Y1	-0.72	0.91	-0.66	0.82	1	1	X5	-0.88	0.83	-0.88	1	1	1
Y2	-0.73	0.91	-0.66	0.82	1	1	Y1	-0.76	0.77	-0.61	0.61	1	1
							Y2	-0.75	0.76	-0.61	0.6	1	1
							Fifth grade						
							X1	X2	X4	X5	Y1	Y2	
X1	1	-0.87	0.79	-0.75	-0.77	-0.79	1	1	0.78	-0.7	-0.82	-0.81	
X2	-0.87	1	-0.71	0.83	0.89	0.88	1	-0.93	1	0.8	0.78	0.78	
X4	0.79	-0.71	1	-0.89	-0.65	-0.66	X2	-0.93	1	-0.81	0.8	0.67	
X5	-0.75	0.83	-0.89	1	0.75	0.75	X4	0.78	-0.81	1	-0.94	-0.67	
Y1	-0.77	0.89	-0.65	0.75	1	1	X5	-0.7	0.8	-0.94	1	0.6	
Y2	-0.79	0.88	-0.66	0.75	1	1	Y1	-0.82	0.78	0.6	0.6	1	
							Y2	-0.81	0.78	-0.67	0.6	1	
							Third grade						
							X1	X2	X4	X5	Y1	Y2	
X1	1	-0.92	0.75	-0.7	-0.69	-0.69	1	1	0.75	-0.7	-0.69	-0.69	
X2	-0.92	1	-0.73	0.79	0.75	0.75	1	-0.92	1	0.79	0.75	0.75	
X4	0.75	-0.73	1	-0.89	-0.58	-0.58	X2	-0.92	1	-0.89	-0.58	-0.58	
X5	-0.7	0.79	-0.89	1	0.65	0.65	X4	0.75	-0.73	1	0.65	0.65	
Y1	-0.69	0.75	-0.58	0.65	1	1	X5	-0.7	0.79	-0.89	1	1	
Y2	-0.69	0.75	-0.58	0.65	1	1	Y1	-0.69	0.75	-0.58	0.65	1	
							Y2	-0.69	0.75	-0.58	0.65	1	

Second grade 7.8% ( $X_1$ ), 0.5% ( $X_2$ ), 5.7% ( $X_4$ ), 0.0% ( $X_5$ )

Third grade 3.0% ( $X_1$ ), 0.0% ( $X_2$ ), 4.5% ( $X_4$ ), 0.5% ( $X_5$ )

Fourth grade 2.5% ( $X_1$ ), 1.1% ( $X_2$ ), 1.4% ( $X_4$ ), 0.7% ( $X_5$ )

Fifth grade 2.2% ( $X_1$ ), 0.7% ( $X_2$ ), 1.1% ( $X_4$ ), 0.0% ( $X_5$ )

These percentages vary greatly not only across the grades but also across the variables. In second grade, for example, 30 students are classified as impaired readers with  $X_1$  and only two with  $X_2$ . Many percentages are far from the expected value (2%). Because of the nonnormality of the variables, the presence of outliers and the level of asymmetry which is different from one grade to another, we define the thresholds for  $Y_1$  and  $Y_2$  in terms of the percentiles. In particular, we choose  $x_{0.5}$  for discriminating a percentage of people higher than the expected percentage. As a matter of fact, the proposed procedure is not intended as a diagnostic test for learning disorders, but it is a screening procedure for detecting students with heavy difficulties in reading. The causes of these difficulties are to be defined by subsequent more detailed analyses.

To investigate the validity of the screening procedure, we may analyse the pairwise correlations between the new variables  $Y_1$  and  $Y_3$  and the variables used in the benchmark tests. The correlation matrices are reported in Table 2. The pairwise correlations are all significantly different from zero ( $\alpha = 0.01$ ). Even though the transformation from  $X_1$  to  $X_2$  and from  $X_3$  to  $X_4$  is not a linear one, the values of these pairs of variables are highly correlated. The transformation from  $Y_1$  to  $Y_2$  is not a perfect linear transformation (because the words have different numbers of syllables), but the correlation is nevertheless equal to 1. The fact that  $Y_1$  and  $Y_2$  are highly correlated with  $X_1$ ,  $X_2$ ,  $X_4$  and  $X_5$  gives evidence that all these variables are a measure of the same aspect of the dyslexia phenomenon.

### Reading Accuracy

In Table 3, we list the values of some univariate statistics of the variables measuring reading accuracy ( $X_3$ ,  $X_6$  and  $Y_3$ ). Figure 3 shows the empirical distributions of these variables through boxplots. As well as the variables measuring the reading speed, variables  $X_3$ ,  $X_6$  and  $Y_3$  have an empirical distribution, which is asymmetric and far from the Gaussian. Whereas the mean and the median values of  $X_3$  and  $X_6$  have a decreasing pattern, from first to fifth grade, the mean and the median values of  $Y_3$  are roughly constant across grades. This different pattern is because the time in the screening procedure is always equal to 1 min, whereas it depends on the ability of the student in the benchmark tests. In the screening, if one student increases the performance from one class to the subsequent class, he/she increases the velocity of reading without penalizing the reading accuracy. Outliers are all in the 'direction of pathology', and this is a desirable property. The normative threshold values for  $X_3$  and  $X_6$  are the 95th percentiles obtained in the study of Sartori et al. (2007). These values are similar to  $x_{0.95}$  obtained in our sample (and reported in Table 3).

Whereas  $X_3$  and  $X_6$  are highly correlated,  $Y_3$  is not correlated with  $X_3$  or with  $X_6$  (Table 4). This is because the time is fixed in the screening procedure. As previously explained, the number of errors has a different pattern from the number of errors in a reading test where the time depends on the ability of the

Table 3. Univariate indices obtained in the sample for variables measuring the reading accuracy

	X <sub>3</sub>					X <sub>6</sub>					Y <sub>3</sub>				
	I	II	III	IV	V	I	II	III	IV	V	I	II	III	IV	V
Min	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Max	84	55	18	19	22	38	31	20	24	27	9	10	6	11	11
X <sub>0.25</sub>	5.0	4.0	2.0	1.0	1.0	4.0	5.0	3.0	3.0	2.0	0.00	0.50	0.50	0.88	0.50
X <sub>0.50</sub>	10.0	6.0	4.0	3.0	2.0	8.0	7.0	6.0	5.0	4.0	1.00	1.50	1.50	1.00	1.00
X <sub>0.75</sub>	17.0	10.3	6.0	5.0	4.0	13.0	10.0	8.0	8.0	6.0	2.00	2.50	2.00	2.50	2.00
X <sub>0.95</sub>	38	19	11	10	9	24	17	14	14	12	4.0	5.0	4.0	4.0	4.0
Mean ( $\bar{x}$ )	13.57	7.96	4.48	3.61	2.8	9.83	8.16	6.29	5.7	4.8	1.30	1.76	1.41	1.69	1.48
S	12.34	6.78	3.56	3.34	3.1	7.20	5.13	4.02	4.4	4.1	1.42	1.51	1.18	1.52	1.53
S/ $\bar{x}$	0.91	0.85	0.79	0.93	1.12	0.73	0.63	0.64	0.76	0.84	1.09	0.86	0.84	0.90	1.03
Skewness	2.07	2.62	1.29	1.63	2.55	1.29	1.18	1.06	1.32	1.92	1.80	1.35	0.97	1.81	2.29
Kurtosis	5.98	11.88	1.91	3.62	8.85	1.72	1.95	1.39	1.97	6.00	4.79	2.97	1.08	5.47	8.63

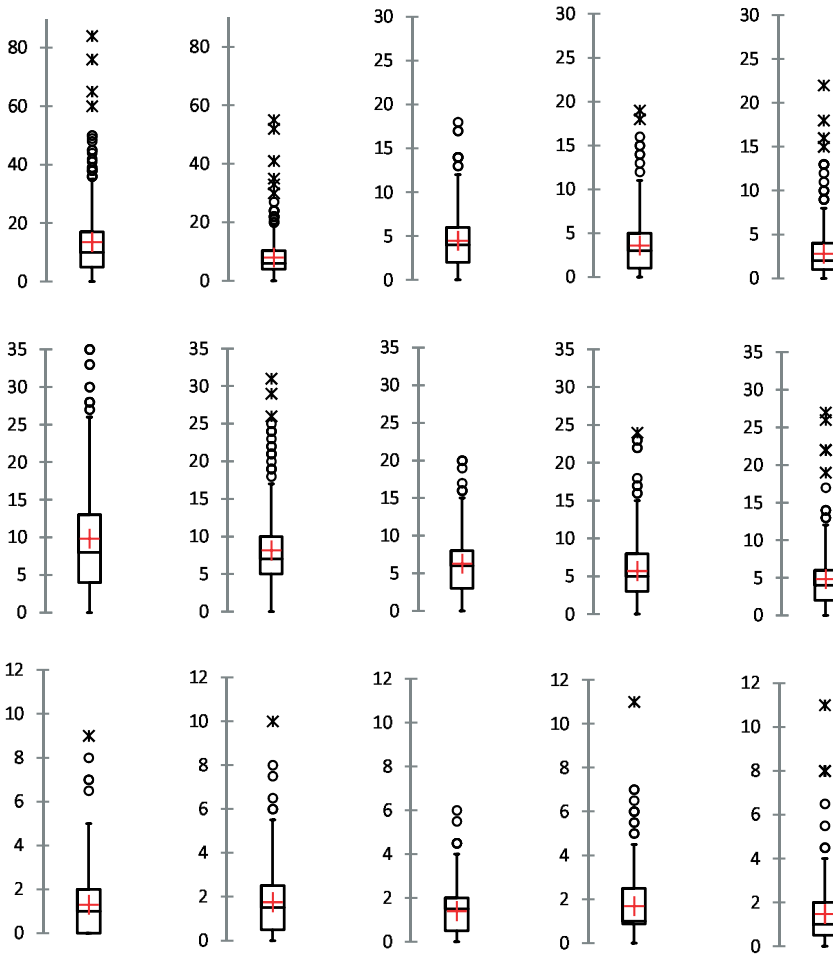


Figure 3. Boxplots of  $X_3$  (first line),  $X_6$  (second line) and  $Y_3$  (third line). First grade is in the first column, second grade is in the second column, third grade is in the third column, fourth grade is in the fourth column and fifth grade is in the fifth column.

Table 4. Correlation matrices for variables measuring reading accuracy

	First grade			Second grade			Third grade			Fourth grade			Fifth grade		
	X3	X6	Y3	X3	X6	Y3	X3	X6	Y3	X3	X6	Y3	X3	X6	Y3
X3	1	0.8	0.37	1	0.69	0.43	1	0.68	0.31	1	0.61	0.34	1	0.61	0.35
X6	0.8	1	0.38	0.69	1	0.39	0.68	1	0.33	0.61	1	0.28	0.61	1	0.3
Y3	0.37	0.38	1	0.43	0.39	1	0.31	0.33	1	0.34	0.28	1	0.35	0.3	1

subject. To further investigate if  $Y_3$  measures the same aspect measured by  $X_3$  and  $X_6$ , we analyse the association between the classifications of the student's performances. Even though two variables are not correlated, they may both lead to the conclusion that a student is a normal reader (or to the conclusion that he/she is an impaired reader). The values of the  $\chi^2$  statistic in the contingency tables obtained

Table 5. Contingency tables

	$Y3 < x_{0.95}$	$Y3 \geq x_{0.95}$	Total		$Y3 < x_{0.95}$	$Y3 \geq x_{0.95}$	Total
$X3 < x_{0.95}$	1327	55	1382	$X6 < x_{0.95}$	1326	57	1383
$X3 \geq x_{0.95}$	65	22	87	$X6 \geq x_{0.95}$	66	20	86
Total	1392	77	1469	Total	1392	77	1469

by considering, for each couple of variables, the category I (the student has a value equal to or greater than  $x_{0.95}$  and is classified as an impaired reader) and the category 0 (the student has a value smaller than  $x_{0.95}$  and is classified as a normal reader) have all a  $p$ -value smaller than 0.05 (Table 5). Therefore, for  $\alpha = 5\%$ , we may reject the null hypothesis of independence between pairs of classifications.

### INTERNAL CONSISTENCY

An explorative factor analysis has been performed in order to investigate the multivariate relationships among variables used in the benchmark tests and in the screening procedure. The analysis, performed on the correlation matrix (obtained with the values of all variables in all grades), shows the presence of two main latent orthogonal factors, both with the principal component (PC) and with the common factors method (CF). The first factor is highly correlated with variables measuring speed and, to a less degree, with  $X_3$  and  $X_6$ . With the PC method, the eigenvalue of this factor is equal to 6 and the percentage of explained variance is 66.7%. With the CF method, the eigenvalue is 5.83 and the percentage of variance is 64.8%. The second factor is highly correlated with  $Y3$  and, to a less degree, with the other variables measuring accuracy ( $X_3$  and  $X_6$ ). With the PC method, the eigenvalue of this factor is 1.28 and the percentage of explained variance is 14.3%. With the CF method, the eigenvalue is 0.86 and the percentage of explained variance is 9.6%. Figure 4 and Table 6 summarize the results obtained

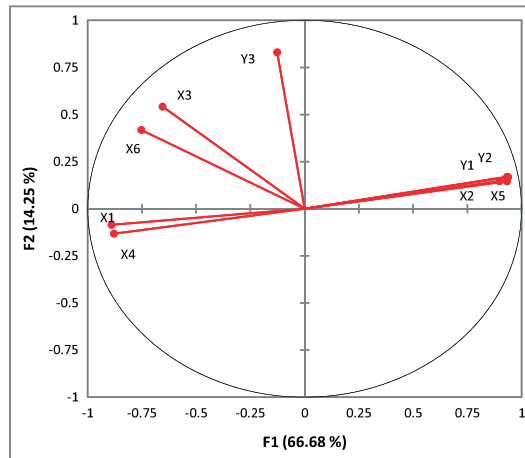


Figure 4. Biplot resulting from factor analysis applied to the correlation matrix. The factors are extracted with the principal component method and are unrotated.

Table 6. Correlations between factors and variables

Factor	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	Y <sub>1</sub>	Y <sub>2</sub>	Y <sub>3</sub>
F <sub>1</sub>	-0.89	0.93	-0.65	-0.88	0.90	-0.75	0.93	0.94	-0.13
F <sub>2</sub>	-0.09	0.15	0.54	-0.13	0.15	0.42	0.17	0.17	0.83

The factors are extracted with the principal component method and are unrotated.

with the PC method. Rotating the factors does not improve the percentage of variance explained by the first two factors. Drawing from these results, with the belief that reading speed and accuracy are two different aspects of reading, we estimate the degree to which the set of variables X<sub>1</sub>, X<sub>2</sub>, X<sub>4</sub>, X<sub>5</sub>, Y<sub>1</sub> and Y<sub>2</sub> measures a single unidimensional latent construct (the reading speed) and the set of variables X<sub>3</sub>, X<sub>6</sub> and Y<sub>3</sub> measures another unidimensional latent construct (the reading accuracy). We estimate the internal consistency of each set of variables by means of the coefficient  $w$  (McDonald, 1999; Zinbarg et al., 2005; 2006), considering the correlation matrix. For the variables regarding speed,  $w = 0.86$ . For the variables regarding accuracy,  $w = 0.64$ . Because these variables have all positive pairwise correlations, we may also calculate the  $\alpha$  coefficient of Cronbach (1951) and their  $\rho^*$  coefficient (Brown, 1910; Spearman, 1910). We obtain  $\alpha = 0.85$  and  $\rho^* = 0.70$ . Regarding the reading speed, if we select variables having positive pairwise correlations (namely X<sub>2</sub>, X<sub>5</sub>, Y<sub>1</sub> and Y<sub>2</sub>),  $\omega$  is 0.94 and both the  $\alpha$  coefficient and the  $\rho^*$  coefficient are 0.98. Even though, in each set, the standardized variables are not  $\tau$ -equivalent and  $\alpha$  gives an overestimation of the internal consistency, all indexes show high inter-correlation among variables belonging to each set.

To evaluate the reliability of the new screening procedure, because this procedure measures two latent construct of the dyslexia phenomenon, we may consider the percentage of variance explained by the first two factors. With both the PC and the CF method of extraction, this percentage is 99.97%, which indicates a very high reliability.

## SYNTHESIS OF THE RESULTS OF THE STATISTICAL ANALYSES

As for the statistical analyses, the following results of this study are relevant both for research and clinical assessment:

- Variables measuring reading speed (both in terms of time or number of words and in terms of syllables read in a second) are not Gaussian, are asymmetric and present many outliers. The skew is always positive in variables measuring the time and the syllables per second in the tests where the time depends on the ability of the reader. For variables measuring the number of words or syllables read in a second in a procedure where the time is fixed, the skew is positive during the first school years and becomes negative during the last school years. The negative skewness and the presence of outliers smaller than  $x_{0.25} - 1.5(x_{0.75} - x_{0.25})$  indicate that these variables have a more discriminative power in the last school years of primary school, when the students read more fluently.

- Variables measuring the number of errors have always a positive skew. This is a desirable property, because the direction of pathology in these variables is positive (impaired readers are students with high values in these variables). Whereas the average number of errors in a test where the time depends on the ability of the students decreases from first to fifth grade, the number of errors in a procedure where the time is fixed has a roughly constant pattern. This means that in a screening with fixed time, if one student increases the performance from one class to the subsequent class, he/she increases the velocity of reading without penalizing the reading accuracy.
- Because of the nonnormality of the distribution, the skewness and the presence of outliers, the threshold values that discriminate between impaired readers and not impaired readers must be estimated in terms of the percentiles, rather than with the mean and the variances. Indeed, the percentiles are more robust to outliers. In our study, the estimates of the means and the variances have been shown to be significantly different from the means and the variances estimated in a previous work, whereas the percentile  $x_{0.95}$  has been shown to remain similar. Moreover, the values  $\bar{x} - 2S$  or  $\bar{x} + 2S$  have been shown to discriminate a percentage of students far from 2%.
- Reading speed and accuracy are two orthogonal latent factors of dyslexia. An explorative factor analysis (performed with different methods of extraction) has highlighted two main orthogonal factors: the first one correlated with the variables measuring speed and the second one correlated with the variables counting errors.
- The set of all variables measuring speed and the set of all variables measuring accuracy have both high internal consistencies. This gives evidence that the number of words and the number of syllables per second read in the screening procedure are a measure of the same feature measured by the variables regarding reading speed in the diagnostic tests. In addition, the counts of errors in the screening procedure are a measure of the same feature measured by the counts of errors on the list of words and pseudo-words.

A limitation of this study is that children with a certified learning disability are not included in the sample. Further studies should include these children in the sample in order to better analyse the discriminating power of the SPILLO screening procedure and of the benchmark tests. Some statistical procedures, such as discriminant analysis and artificial neural networks, aimed at determining the set of variables that best separate between groups (in our case, between impaired and not impaired readers), require the knowledge of the group membership of observations in the sample. Administering the screening procedure and the word and pseudo-words reading tests to children with a certified learning disability may also allow estimating the probability of a correct diagnosis and the probability of false positive.

## CONCLUSIONS

The aim of this study was to develop a new screening tool to identify children at risk of reading difficulties in primary school. Results and analysis confirm that SPILLO is a good instrument to be used as a screening test, as it is able to discriminate impaired readers among normal readers: the reliability of this procedure, as measured by the percentage of variance explained by the first two factors, is 99%.



Screening is a methodology used to indicate, with a good level of precision, whether the subjects tested are at risk in a certain skill. However, a screening result is not comparable with the result of an accurate diagnostic assessment. The aim of a screening is not to formulate diagnoses but to identify a population at risk, who should then participate in a more specific assessment, which has a good correlation in finding people who need a diagnosis and excluding people who do not need any further assessment (Wood, Flowers, Meyers & Hill, 2002; Choi, 1982; 1992; 1996). In order to meet the previously mentioned criteria, (sensitivity and specificity), the text used for our screening task was selected on the basis of the average reading skills of primary school students. In order to set up a sensitive task, the readability of the text was chosen at a difficult level (score 71), but it contains a high percentage (95.57%) of words found in the VdB. The same text was used for all grades, and this choice has some advantages: (i) it allows a measure of reading skill improvement across ages and grades; (ii) it reduces methodological errors in screening made by examiners; and (iii) Italian studies (Tressoldi, 2008) demonstrated that, although reading speed and accuracy are influenced by semantic and orthographic features, these effects are stable across ages. Statistical analysis indicated that a 1-min reading test is suitable to identify, with a good degree of precision and certainty, subjects who are at risk of developing reading difficulties.

Moreover, our findings are consistent with previous studies in transparent orthographies reporting that reading disorder is more evident in terms of speed rather than of accuracy (Wimmer, 1993). Sometimes the reading can be accurate but is typically slow and stunted. In our sample, the percentage of variance explained by the speed factor is 66.7%, whereas the percentage of variance explained by the accuracy factor is 14.3%

Zoccolotti, De Luca, Di Pace, Judica, Orlandi and Spinelli (1999), Zoccolotti, De Luca, Judica and Spinelli (2007) and Zoccolotti and Burani (2010) showed that children with a reading disorder can make a different number of mistakes but proceed in the written text with tiredness and difficulties. In addition, studies on the development of reading in Italian dyslexics (Campanini, Battafarano & Iozzino, 2010; Re, Tressoldi, Cornoldi & Lucangeli, 2011; Tressoldi, 1996; Tressoldi, Stella & Faggella, 2001; Tressoldi, Vio & Iozzino, 2007; Tressoldi, Lorusso, Brenbati & Donini, 2008) confirm that the evolution of the disorder is faster for accuracy rather than for speed, which evolves more slowly. Speed seems to be the most reliable prognostic indicator for predicting the subsequent development of reading skills.

Finally, some general remarks on social relevance, as a recent Italian act of parliament on learning disorder assigned this new duty to schools and recommended primary school children should be screened for reading difficulties by their teachers under the supervision of professionals in the field. These procedures must foster students' development of skills involved in school learning (AID, 2009). The use of SPILLO, a computer-based test, gives advantages to examiner in recording and scoring and makes this instrument an effective screening tool, which can also be used to test large populations without any specific training. The use of software that accesses the electronic database allows a more objective evaluation, as well as the possibility to rapidly obtain a large amount of data on reading skills at different scholastic levels. Finally, teachers could decide to use this test to check the average reading skills of their class or to determine the level of some of their students with the aim of adopting teaching strategies based on their

students' abilities. Given its feasibility, the test can be periodically used to monitor each student's progress and to indicate the necessity for students who are at risk to be given extra help. Through the administration of this test after a given period, it is possible to verify whether a child continues to read at a level that requires specific intervention (and therefore he/she is likely to have learning disorders), or whether the child has overcome the difficulty.

## REFERENCES

- AID Associazione Italiana Dislessia, Comitato Promotore Consensus Conference (Eds.) (2009). *Disturbi Evolutivi Specifici di Apprendimento Raccomandazioni per la pratica clinica di dislessia, disortografia, disgrafia e discalculia*. Erickson, Trento, Italy.
- Allemandri, V. (2005). Screening per l'individuazione precoce dei bambini a rischio di dislessia. *Dislessia*, 2, 209–228.
- Angelelli, P., Judica, A., Spinelli, D., Zoccolotti, P., & Luzzatti, C. (2004). Characteristics of writing disorders in Italian dyslexic children. *Cognitive & Behavioral Neurology*, 17, 18–31.
- Badian, N. A. (1982). The prediction of good and poor reading before kindergarten entry: a 4-year follow-up. *The Journal of Special Education*, 16, 309–318.
- Badian, N. A. (2000). Do preschool orthographic skills contribute to prediction of reading? In Badian N. (Eds.) *Prediction and prevention of reading failure*. Timonium, MD: York Press, 31–56.
- Bellocchi, S. (2011). Prove ZERO: un protocollo per l'identificazione precoce di difficoltà di letto-scrittura. *Dislessia*, 8, 81–93.
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3, 296–322.
- Campanini, S., Battafarano, R., & Iozzino, R. (2010). Evoluzione naturale della lettura del brano, delle liste di parole e non parole e della comprensione del testo in dislessici mai trattati. *Dislessia*, 7, 165–179.
- Cornoldi, C., & Colpo, G. (1986). Prove di rapidità e correttezza nella lettura del gruppo MT. Organizzazioni Speciali, Firenze, Italy.
- Cornoldi, C., & Colpo, G. (1998). Prove di lettura MT per la scuola Elementare-2. Organizzazioni Speciali, Firenze, Italy.
- Cossu, G. (1999). The acquisition of Italian orthography. In M. Harris, G. Hatano (Eds.) *Learning to read and write: a cross-linguistic perspective*. Cambridge, UK: Cambridge U.P.
- Choi, B. C. K. (1982). Index for rating predictive accuracy of screening tests. *Methods of Information in Medicine*, 21, 149–153.
- Choi B. C. K. (1992). Sensitivity and specificity of a single diagnostic test in the presence of work-up bias. *Chinese Journal of Epidemiology*, 45, 581–586.
- Choi, B. C. K. (1996). Causal modeling to estimate sensitivity and specificity of a test when prevalence changes. *Chinese Journal of Epidemiology*, 8, 80–85.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- De Beni, R., Cisotto, L., & Carretti, B. (2001). *Psicologia della lettura e della scrittura*. Erickson, Trento, Italy.
- De Mauro, T. (1997). *Guida all'uso delle parole*. Editori Riuniti, Roma, Italy.
- Faglioni, P., Gatti, B., Paganoni, A. M., & Robutti, A. (1967). La valutazione psicometrica della dislessia. *Infanzia anormale*, 81, 628–661.
- Flesch, R. F. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221–223.
- Franchina, V., & Vacca, R. (1986). Taratura dell'indice di Flesch su testo bilingue italoinglese di unico autore. *Atti dell'incontro di studio su Leggibilità e Comprensione*. Coop. Spazio Linguistico, Roma, Italy, 47–49.
- Goswami, U., Gombert, J. E., & Fraca de Barrera, L. (1998). Children's orthographic representation and linguistic transparency: nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics*, 19, 19–52.

- Hermann, J. A., Matyas, M., & Pratt, C. (2006). Meta-analysis of the nonword reading deficit in specific reading disorder. *Dyslexia*, 12(3), 195–221.
- Holliman, A. J., Wood, C., & Sheehy, K. (2010). The contribution of sensitivity to speech rhythm and non speech rhythm to early reading development. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 30, 247–267.
- Invernizzi, M., Justice, L., Landrum, T. J., & Booker, K. (2004). Early literacy screening in kindergarten: widespread implementation in Virginia. *Journal of Literacy Research*, 36, 479–500.
- Lastrego, C., & Testa, F. (1995). *Ho paura*. Editrice Piccoli, Torino, Italy.
- Jenkins, J. R., & O'Connor, R. (2002). Early identification and intervention for young children with reading/learning disabilities. In R. Bradley, L. Danielson, D. Hallahan (Eds.) *Identification of learning disabilities* (pp. 99–149). Hillsdale, NJ: Erlbaum.
- Justice, L. M., Invernizzi, M. A., & Meier, J. D. (2002). Designing and implementing an early literacy screening protocol: suggestions for the speech-language pathologist. *Language, Speech, and Hearing Services in Schools*, 33, 84–101.
- McDonald, R. P. (1999). *Test theory: a unified treatment*. Erlbaum, 90103.
- Nicolson, I. R., & Fawcett, A. J. (1997). Development of objective procedures for screening and assessment of dyslexic students in higher education. *Journal of Research in Reading*, 20, 77–83.
- O'Connor, R. E., & Jenkins, J. R. (1999). The prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading*, 3, 159–197.
- Orsolini, M., Fanari, R., & Maronato, C. (2005). *Difficoltà di lettura nei bambini*. Carocci, Roma, Italy.
- Piemontese, M., & Cavaliere, L. (1997). Leggibilità e comprensibilità dei sussidiari per le scuole elementari. In R. Cal, S. Ferreri (Eds.) *Il testo fa scuola. Libri di testo, linguaggi ed educazione linguistica*. Atti dell'VIII Convegno nazionale Giscel (pp. 21–23). Palermo: La Nuova Italia Marzo, Firenze, Italy.
- Ransby, M. J., & Swanson, H. L. (2003). Reading comprehension skills of young adults with childhood diagnoses of dyslexia. *Journal of Learning Disabilities*, 36, 538–555.
- Re, A. M., Tressoldi, P. E., Cornoldi, C., & Lucangeli, D. (2011). Which tasks best discriminate between dyslexic university students and controls in a transparent language? *Dyslexia*, 17, 227–241.
- Rathvon, N. (2004). *Early reading assessment: a handbook for practitioners*. New York: Guilford Press.
- Riccardi Ripamonti, I., Cividati, B., Truzoli, R., & Russo, V. (2007). Evoluzione delle modalità di lettura nel primo anno della scuola primaria e metodi di insegnamento. *Dislessia*, 4, 13–26.
- Sartori, G., Job, R., & Tressoldi, P. E. (1995). *DDE Batteria per la valutazione della dislessia e della disortografia evolutiva*. Firenze, Italy: Giunti O.S.
- Sartori, G., Job, R., & Tressoldi, P. E. (2007). *DDE-2 Batteria per la valutazione della dislessia e della disortografia evolutiva - 2*. Firenze, Italy: Giunti O.S.
- Scarborough, H. S. (1998). Predicting the future achievement of second graders with reading disabilities: contributions of phonemic awareness, verbal memory, rapid serial naming, and IQ. *Annals of Dyslexia*, 48, 115–136.
- Schulte-Krue, G. (2001). *Lese-Rechtschreibstrung und Sprachwahrnehmung-Psychometrische und neurophysiologische Untersuchungen zur Legasthenie. (Dyslexia and Speech Perception)*. Mnster, Germany: Waxman Verlag.
- Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143–174.
- Snow, C., Burns, M. S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children, committee on the prevention of reading difficulties in young children, commission on behavioural and social sciences and education, National Research Council*. Washington, DC: National Academy Press.
- Spearman, C. C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3, 271–295.
- Stella, G. (2004). *La dislessia. Quando un bambino non riesce a leggere: cosa fare, come aiutarlo*. Bologna, Italy: Il Mulino.
- Stella, G., & Apolito, A. (2004). Lo screening precoce nella scuola elementare. Può una prova di 16 parole prevedere i disturbi specifici di apprendimento? *Dislessia*, 1, 111–118.

- Tressoldi, P. E. (1996). L'evoluzione della lettura e della scrittura dalla scuola elementare alla terza media. Dati per un modello di apprendimento e per la diagnosi dei disturbi specifici. *Età evolutiva*, 56, 43–55.
- Tressoldi, P. E. (2008). I brani della batteria MT si possono leggere tutti con la stessa velocità? Norme trasversali dal secondo anno della primaria al terzo della secondaria di primo grado. *Dislessia*, 5, 339–345.
- Tressoldi, P. E., Stella, G., & Faggella, M. (2001). The development of reading speed in Italians with dyslexia: a longitudinal study. *Journal of Learning Disabilities*, 34, 414–417.
- Tressoldi, P. E., Vio, C., & Iozzino, R. (2007). Efficacy of an intervention to improve fluency in children with developmental dyslexia in a regular orthography. *Journal of Learning Disabilities*, 40, 203–209.
- Tressoldi, P. E., Lorusso, M. L., Brenbati, F., & Donini, R. (2008). Fluency remediation in dyslexic children: does age make a difference? *Dyslexia*, 14, 142–152.
- Willburger, E., & Landerl, K. (2009). Der Ein-Minuten Leseflüssigkeitstest-ein Verfahren zur Diagnose der Leistung im Wort- und Pseudowortlesen [1-Minute Reading Fluency Test]. *Test und Trends*, 7, 65–80.
- Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics*, 14, 1–33.
- Wood, C., & Terrell, C. (1998). Poor readers' ability to detect rhythm and perceive rapid speech. *British Journal of Developmental Psychology*, 16, 397–413.
- Wood, F., Flowers, L., Meyers, M., & Hill, D. (2002). How to evaluate and compare screening test: principles of science and good sense. IDA 35rd Annual Conference, Atlanta, Georgia.
- Zinbarg, R., Revelle, W., Yovel, I., & Li, W. (2005). Cronbachs, Revelles, and McDonalds: their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70, 123–133.
- Zinbarg, R., Yovel, I., Revelle, W., & McDonald, R. (2006). Estimating generalizability to a universe of indicators that all have an attribute in common: a comparison of estimators for *wh*. *Applied Psychological Measurement*, 30, 121–144.
- Zoccolotti, P., De Luca, M., Di Pace, E., Judica, A., Orlandi, M., & Spinelli, D. (1999). Markers of developmental surface dyslexia in a language (Italian) with high grapheme-phoneme correspondence. *Applied Psycholinguistic*, 20, 191–216.
- Zoccolotti, P., De Luca, M., Judica, A., & Spinelli, D. (2007). Un progetto di studio e di intervento sui disturbi specifici dell'apprendimento della lettura nei bambini italiani. *Neuropsicologia evolutiva e disturbi dell'apprendimento*, 33, 23–40.
- Zoccolotti, P., De Luca, M., Di Filippo, G., Judica, A., & Martelli, M. (2008). Reading development in an orthographically regular language: effects of length, frequency, lexicality and global processing ability. *Reading and Writing: An Interdisciplinary Journal*, 22, 1053–1079.
- Zoccolotti, P., & Burani, C. (2010). Apprendimento e disturbi di lettura in ortografie trasparenti. In Vicari S., Caselli M. C. (Eds.) *Neuropsicologia dello sviluppo*. Il Mulino, Bologna, Italy: 179–195.

Copyright of Dyslexia (10769242) is the property of John Wiley & Sons, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.