Received: 13 October 2022 | Accepted: 15 March 2023

DOI: 10.1111/opo.13134

ORIGINAL ARTICLE

OPO THE COLLEGE OF OPTOMETRISTS

New Italian version of the Wilkins Rate of Reading Test: Materials for repeated-measure designs in optometry and neuropsychological research

Fabrizio Zeri[1,2,3] | **Silvia Tavazzi**[1,2] | **Marco Punzi**[1] | **Federica Miglio**[1,2] | **Bruce J. W. Evans**[4] | **Maria De Luca**[5]

[1]Department of Materials Science, University of Milano-Bicocca, Milan, Italy

[2]COMiB Research Centre in Optics and Optometry, University of Milano-Bicocca, Milan, Italy

[3]College of Health and Life Sciences, Aston University, Birmingham, UK

[4]The Institute of Optometry, London, UK

[5]IRCCS Fondazione Santa Lucia, Rome, Italy

Correspondence
Silvia Tavazzi, Department of Materials Science, University of Milano-Bicocca, Milan, Italy.
Email: silvia.tavazzi@unimib.it

Funding information
Italian Ministry of Health (Ministero della Salute)

Abstract

Purpose: To develop a series of equivalent passages of text in Italian, according to the principles of the Wilkins Rate of Reading Test (WRRT), suitable for both clinical examination and scientific research when equivalent stimuli are needed to compare performance in repeated-measure designs.

Method: Fifteen high-frequency Italian words (matched for grammatical class and length to the English WRRT) were used to generate 15 different 10-line meaningless passages, according to the design principles of the English WRRT. Thirty-two healthy Italian-speaking higher education students read the passages aloud according to a fixed randomisation schedule. Performance was recorded digitally to measure reading speed and accuracy offline. Equivalence between the passages and the practice and fatigue effects for both reading speed and accuracy were examined as well as test–retest reliability.

Results: No significant difference in reading speed and accuracy was found between the passages. There was a significant practice effect on reading speed but not accuracy, with the first presented passage read significantly slower than the others. There was no evidence of a fatigue effect. Reading speed, the reference metric for the WRRT, showed good test–retest reliability.

Conclusions: The passages of the Italian version of the WRRT were equivalent to each other. The practice effect suggests that familiarisation with the test (i.e., reading at least one matrix of words) should be carried out before consecutive/repeated reading of different passages for experimental or clinical purposes.

KEYWORDS
accuracy, reading, reading speed, Wilkins Rate of Reading Test

INTRODUCTION

The Rate of Reading Test (RRT), most commonly known as the Wilkins RRT (WRRT) from the name of its designer, has been developed and used in the field of optometry to assess visual performance in a reading-like task.[1–5] The test is designed to be visually demanding but requires only very basic reading skills, so performance is minimally affected by language skills. Therefore, the test is useful to study the effect of visual factors and interventions on reading.

The original test comprised two separate passages,[3] soon replaced with four,[6] with each made up of 10 lines. Each line was formed by the same 15 words arranged in a pseudo-random order.[3] Wilkins used sequences of unrelated words to isolate visual input to reading with minimal requirements for higher cognitive processing related to characteristics of

language such as syntax or semantics. Typically, the user undertakes the test four times (each time with a different passage), twice with the intervention under investigation (A) and twice with no intervention or with a control intervention (B) in the order ABBA. Each passage is scored as words correctly read per minute (wpm) and the two scores with the intervention are averaged and compared to the mean of the two scores without the intervention.

In the recent study by Gilchrist,[6] the reading speed with the WRRT for English readers was quicker in adults (average of 159 and 184 wpm in the two samples of adults analysed) than in children (average of 95 and 111 wpm in the two samples of children analysed) with a precision (within-subjects standard deviation [SD], s_w) that varied from 3.8 wpm to 11.5 wpm among the samples of different ages. Moreover, WRRT showed a very high test–retest reliability with an intraclass correlation coefficient or Pearson correlation coefficient between the first and second tests >0.80.[6–8]

Wilkins' website (https://www1.essex.ac.uk/psychology/overlays/rrt%20OC4.htm) provides the test in other languages (German, Italian, Portuguese and Arabic), and the authors are aware of additional non-English versions by researchers in other countries/languages, including British Sign Language,[9] modern Hebrew, Norwegian and Dutch. However, a PubMed search reveals a validation study of only one non-English version, namely Kannada.[10] Although other non-English versions have been used in studies[11,12] to evaluate the effect of coloured overlays on reading performance, their psychometric properties do not appear to have been investigated further.

Therefore, the authors FZ, ST, MP and MDL developed a new Italian version of the WRRT as part of a suite of psychophysical and psychometric tools to evaluate reading performance in clinical practice and research[13,14] such as the MNREAD charts,[15] the Radner test[16] and the NAVQ questionnaire.[17]

Wilkins (personal communication) has highlighted that versions of WRRT in different languages should not be a direct translation, but a transliteration whereby common words, preferably those that are monosyllabic and within the 100 most frequently used words in that language are presented in random order, with the constraint that two identical words do not appear one above the other on neighbouring lines, and that the last word in each line differs from the first word in the next line. The first Italian version of the WRRT (https://www1.essex.ac.uk/psychology/overlays/rrt%20OC4.htm) comprises two passages on separate pages. This version uses longer words than the English one, both in terms of characters per word and syllables per word (e.g., 53% of the words chosen for the Italian version are five- or six-letter disyllabic words, whereas in the English version, there were no five- or six-letter words, and all words had one syllable). Therefore, compared with the English version, the first Italian version has longer lines (62 instead of 44 characters, excluding blanks), potentially adding visual complexity that has not been addressed formally.

The aim of the present study was to develop and validate a new series of equivalent WRRT passages in Italian, according to the main principles of the WRRT.[3] It was decided to create 15 different passages of text, which is far more than the usual four that are used clinically. This is to facilitate repeated-measure designs, when clinical examination or scientific research requires measuring performance under multiple different conditions or many successive repetitions of the same condition (e.g., in the fields of optometry, education or psychology).

Along with equivalence between the passages, the effects of practice and fatigue were examined. Like the original English version, reading speed and accuracy were not meant to examine conventional reading skills, but to evaluate performance in clinical or experimental tasks (e.g., to compare in a within-subject design different optical corrections, coloured filters, etc.). The equivalent passages can also be conveniently adopted in the field of neuropsychology, when it is necessary to collect measurements of many successive repetitions of the same condition or measure the effect of multiple visual-perceptual manipulations (e.g., when studying the reading process in patients with visual field defects or attentional deficits) on performance in a reading-like task.

The present study was intended to make available an Italian series of WRRT passages that were equivalent, and to investigate practice and fatigue effects in a sample of healthy young volunteers.

Key points

- The Wilkins Rate of Reading Test (WRRT) measures the effect of visual factors on reading; therefore, several versions of the test could be useful in optometry, education or neuropsychology when repeated measurements are required.
- A significant practice effect limited to the first passage was found. In case of repeated testing, it is important to allow practice with one passage before starting measurements.
- The large number (15) of parallel versions of this Italian WRRT was shown to be equivalent to one another and have good test–retest reliability.

METHODS

Development of the Italian WRRT passages

Fifteen Italian words were chosen with characteristics of length (four 2-letter, eight 3-letter and three 4-letter words), and grammar class (one adverb, one conjunction, one definite article, two pronouns, two nouns, three prepositions and five verbs), consistent with the WRRT. The 15 words were high-frequency words belonging to the so-called fundamental words of the Italian language.[18] There were nine monosyllabic and six disyllabic words. The words were

di [of], *ha* [has], *si* [third person reflexive pronoun, used in reflexive verbs], *la* [the], *amo* [I love], *che* [that/which], *con* [with], *era* [was], *fai* [do], *non* [not], *per* [for], *una* [one], *anno* [year], *sono* [am/are] and *uomo* [man].

In the English WRRT,[3] a passage of the test consists of 10 lines, with 15 words ordered differently on each line. Rules were applied so that each word occurred only once in a line and once in a column, and the last word in a line was different from the first word in the next line. Since the aim of the present study was to make available a ready-to-print set of many equivalent passages, a total of 15 passages (labelled with consecutive letters of the alphabet from A to O) were developed following the rules described above. Across passages, an additional rule was introduced, so that a given sequence of 15 words appeared only once in a line. Finally, every line of 15 words was individually checked to avoid the casual presence of syntactic or semantic associations between words: in case of such an occurrence, the line was changed.

The typeset adhered to the typographic specifications of Wilkins,[3,6] that is Times New Roman 9-point font, single-spaced lines with 4-point horizontal spacing between words with an interline space of 3.15 mm. The final layout of the passage was set as a paragraph 72.5 mm wide and 33.4 mm high. Each of the 15 passages was arranged in a separate page of a Microsoft Word file (Microsoft.com) and printed at 1200 dpi resolution.

Pilot study

Twenty-four children attending the beginning of second grade (mean age = 7.7 years, SD = 0.3, range = 7.2–8.4) participated in a brief pilot study designed to verify that the set of words proposed to make up the final passages could easily be read by individuals at an early stage of formal education. Participation of the children was approved by their parents who provided written informed consent. The children read the 15 words aloud (see the next paragraph) presented one at a time on the liquid-crystal display (LCD) monitor of a computer through Microsoft Powerpoint (Microsoft.com). The words were presented in Times New Roman, 80-point font, chosen to be supra-threshold to avoid any visual problem (an x-height of 1.3 cm at a viewing distance of 40 cm, or equivalent to about 1.35 logMAR). None of the children made outright reading errors. Only two children hesitated as they pronounced the word '*uomo*' (man) (i.e., they made a brief silent pause after having pronounced the first vowel). All the children reported there were no unfamiliar words and showed understanding of their meanings.

Participants

Participants of the main study were recruited from students of the University of Milano-Bicocca (Milan) according to the following inclusion criteria: no ocular pathology; no known reading impairment; monocular best-corrected distance visual acuity (BCVA) at distance ≤0.10 logMAR in each eye; no significant ocular motility or binocular vision anomalies (including strabismus); near point of convergence ≤10 cm; stereoscopic acuity ≤80 arcsec; binocular amplitude of accommodation ≥8 D and binocular accommodative facility with ±2.00 D lenses ≥5 cpm. These inclusion criteria were selected to exclude participants whose visual status could make them prone to fatigue, since a goal of the study was to investigate the effect of normal fatigue on performance with the WRRT. All volunteers provided written informed consent.

Thirty-two participants (18 males and 14 females; mean age = 22.9 years, SD = 1.8, range = 20.1–27.2; mean number of years of formal education = 15.7 years, SD = 0.9, range 14–17 years) participated in the study. All were invited to return for retesting after 2 weeks and 19 subjects agreed (8 males and 11 females; mean age = 22.9 years, SD = 1.9, range = 20.1–26.5; mean number of years of formal education = 15.8 years, SD = 0.9, range 14–17 years).

The chosen sample size complied with the need to verify the difference between the means of two repeated measures (test–retest) of reading speed (wpm). G*Power software (gpower.hhu.de/) was used to determine the sample size. The analysis was based on matched pairs *t*-test (two-sided). An effect size of 0.76 was calculated for a difference between two measurements of 10 wpm (which is a significant clinical difference) considering means, SDs and a correlation between repeated measures already available from our laboratory. Using an α error and $1-\beta$ (power) of 0.05 and 0.80, respectively, the required sample size was $N = 18$.

Procedure

All procedures were undertaken in accordance with the Declaration of Helsinki and approved by the Board of Optics and Optometry of the University of Milano-Bicocca (9 September 2019).

Visual assessment

A comprehensive eye and visual examination was performed on each participant at the Research Centre in Optics and Optometry at the University of Milano Bicocca. Direct ophthalmoscopy and slit-lamp biomicroscopy were carried out to detect any ocular anomaly. Monocular non-cycloplegic subjective refraction was undertaken with a phoropter procedure, followed by a final binocular equalisation with dissociated testing to obtain the least minus/most plus correction compatible with good visual acuity.

High-contrast monocular BCVA was measured at 5 m using Sloan letters displayed in five-letter lines with the size decreasing logarithmically according to the principle of the

ETDRS chart. Letters were displayed on an LCD optotype system (Vision Chart, CSO; csoitalia.it/) at high-contrast (97.8 ± 0.2%), when measured with a photometer (Chroma Meter CS-100A, Minolta; konicaminolta.eu/eu-en/navigation/featured-business-areas/measuring-instruments) under photopic conditions (450 ± 50 lux, measured by a luxmeter HT307, HT Italia; ht-instruments.it/it-it/). A forced-choice procedure and a letter-by-letter (0.02 logMAR) scoring criteria were used to assess the threshold in logMAR units. Ocular motility was examined using an H pattern test.

The tests described below were all undertaken with the distance subjective refraction placed before the eyes in a trial frame. Binocular amplitude of accommodation was measured by the Donders' push-up method using a Royal Air Force rule (Bernell Corporation; bernell.com) and N5 stimulus, moved at constant rate of approximately 1 cm/s, starting from 30 cm.[19] The target positions representing both first blur and total blur were recorded. In addition, binocular accommodation facility at 40 cm with ±2.00 D lenses was tested for 60 s. A Bernell n.9 vectogram (bernell.com) was used to check for suppression, and participants were asked to report when the 6/9 target appeared clear. The near point of convergence test was performed by slowly moving an accommodative target (a single letter of approximately 0.2 logMAR equivalent at 40 cm) towards the eyes until the subject either reported diplopia or the examiner noticed a break in fusion; the recovery position was also assessed. Stereoacuity was measured at 40 cm with the Wirt circles of the FLY Stereo Acuity Test (Vision Assessment Corporation; visionassessment.com/). The visual assessment was undertaken in a separate session/day before the reading session.

Wilkins Rate of Reading Test

Each page containing a single passage was displayed on a reading desk at a viewing distance of 40 cm (layout characteristics of the passage are described above). Participants read the passages under photopic conditions (550 ± 50 lux, measured by the HT307 luxmeter) with an average luminance of the paper surface (eight measurements around the text) of 135 ± 11 cd/m^2 (Chroma meter CS 100 A; Minolta; konicaminolta.eu/eu-en/navigation/featured-business-areas/measuring-instruments). All participants who normally used a refractive correction (spectacles or contact lenses) for their academic work wore this correction during the reading session (21 out of 32); otherwise, no refractive correction was provided. The difference between the habitual and the subjective correction evaluated during the visual assessment was nil in 11 cases and negligible in the remaining 10 cases, that is, unlikely to interfere with the reading task (see the Results section for details). In the remaining 11 participants with no habitual correction, the reading task was performed without optical correction because seven were emmetropic, and the remaining four had only a negligible subjective correction.

Participants were asked to read the entire passage aloud as quickly and accurately as possible. An interval of 1 min was allowed between passages. The reading was recorded digitally. The order of the passages from A to O was randomised across participants using six different reading sequences.

Reading speed (wpm) based on words correctly read, and accuracy (percentage of reading errors) were monitored during the experiment and measured offline by replaying the recordings afterwards.

Reading errors were scored as word substitution (replacing the word with another) or word omission (skipping of a word), omitting an entire line (scored as one error), insertions (repeating the previously uttered word or inserting another word) and production of a non-word (a pronounceable string of letters that is not in the lexicon). All the tests and retests were carried out in the same way following detailed written instructions which were read to the participants. For each participant, test and retest were carried out by the same examiner, who was masked to the results of the first test.

Data analysis

Data analysis was carried out using SPSS software (ibm.com). Descriptive statistics were used to illustrate the demographic characteristics of the participants, the results of the visual assessment and the reading performance, that is reading speed (wpm) and reading accuracy (percentage of errors) averaged separately for each passage from A to O.

The Shapiro–Wilk normality test was used to determine if reading speed and accuracy followed a normal distribution. ANOVA was used for normally distributed variables whereas the Friedman test was used for non-parametric analyses. Repeated-measures analyses were run to (1) examine the equivalence between the 15 passages (from A to O) and (2) investigate any practice and fatigue effects that may have been due to the order of the readings. Bonferroni and Dunn–Bonferroni corrections were used to correct for multiple comparisons, for parametric and non-parametric post-hoc analyses, respectively. Test–retest reliability was evaluated for each of the 15 passages (from A to O) by the intraclass correlation coefficient (ICC) based on single measure, consistency type and two-way mixed effects model,[20] and 95% confidence intervals (CIs) were calculated. A comparison between test and retest for reading speed and accuracy was performed by a matched-pairs test (*t*-test, or Wilcoxon depending on the normality of the data distribution). Bland–Altman plots were used to assess the difference in test–retest measurements as a function of the mean of the two measurements.[21] The presence of a proportional bias in the Bland–Altman plots was explored by examining the correlation between the average and the difference between the two measurements (i.e., test and retest).

RESULTS

Visual assessment

For all participants, the average difference between habitual correction and the subjective correction assessed by the procedure described in the 'Visual Assessment' paragraph was negligible; the spherical equivalent refractive error was slightly more negative when measured subjectively when compared with the habitual correction (−0.15 and −0.12 D in the right and left eyes, respectively; range −0.75 to +0.13 D). The optometric profile of participants evaluated in the visual assessment procedure all showed acceptable visual function in terms of visual acuity, accommodation and binocular vision. Descriptive statistics for the main optometric variables are provided in Table 1.

TABLE 1 Visual assessment.

Variable	Mean	SD	Min	Max
OD SER (D) of subjective refraction	−1.93	2.13	−5.75	1.25
OS SER (D) of subjective refraction	−1.93	2.12	−6.50	1.00
OD BCVA (logMAR)	−0.10	0.07	−0.26	0.00
OS BCVA (logMAR)	−0.08	0.07	−0.24	−0.24
NPC break (cm)	5.6	3.5	0.0	10.0
NPC recovery (cm)	8.5	4.1	2.0	15.0
NPA first blur (D)	12.9	4.1	7.4	25.0
NPA total blur (D)	21.1	6.9	10.0	33.3
Accommodation facility (cycles per min)	10.5	4.6	5.0	21.0
Stereoacuity (arcsec)	22.5	4.3	20	40

Note: Main optometric characteristics of participants.

Abbreviations: Arcsec, arc seconds; BCVA, best corrected visual acuity; D, dioptres; logMAR, logarithm of the minimum angle of resolution; NPA, near point of accommodation; NPC, near point of convergence; OD, right eye; OS, left eye; SD, standard deviation; SER, spherical equivalent refraction.

Reading speed and accuracy

Equivalence between the different passages

Descriptive statistics are reported in Table 2. Figure 1 illustrates reading speed and accuracy for the 15 passages. Reading speed was normally distributed for all of the passages (Shapiro–Wilk test: $p > 0.05$). Reading accuracy was not normally distributed (Shapiro–Wilk test: all $p_s < 0.05$). The repeated-measures ANOVA did not show any significant difference in reading speed between the 15 passages (F (1, 14) = 0.90; $p = 0.53$). Regarding reading accuracy, the Friedman test did not show any significant difference between the various passages ($r = 12.1$, $p = 0.60$). Passage F showed a high mean and SD for reading accuracy because a single participant made many repetitions and inversions of words during his first attempt. The error rate of this participant decreased after the first reading. Overall, the passages were read at an average speed between 165 and 170 wpm, with an error rate between 2.1% and 3.0%.

TABLE 2 Equivalence between the passages.

Passage name	Reading speed (wpm)				Accuracy (percentage of errors)				
	Mean	SD	Min	Max	Mean	Median	SD	Min	Max
A	168.2	24.7	123.5	221.2	2.7	2.3	2.5	0	13.3
B	168.1	22.3	119.5	217.2	2.9	2.3	2.4	0	8.5
C	169.0	23.3	117.6	217.7	2.2	2.0	1.6	0	6.8
D	169.7	21.6	115.2	207.7	2.7	2.0	2.3	0	8.5
E	167.2	22.7	120.5	214.8	2.8	2.3	1.9	0	7.3
F	165.6	25.0	113.3	217.8	3.0	2.0	3.7	0	19.8
G	169.1	27.3	119.5	232.5	2.2	1.7	1.7	0	5.4
H	164.9	22.8	125.0	205.9	2.5	2.3	1.8	0	7.7
I	169.6	25.1	125.1	217.3	2.3	2.0	1.8	0	5.8
J	165.7	25.0	114.0	223.5	2.2	1.3	1.8	0	6.7
K	168.2	24.6	122.5	220.3	2.4	1.7	2.2	0	7.7
L	166.8	22.4	118.2	214.2	2.4	2.0	2.2	0	9.8
M	165.7	25.2	115.5	220.7	2.8	2.3	2.2	0	7.7
N	165.2	24.1	116.0	223.1	2.1	1.3	2.0	0	8.2
O	166.6	27.5	118.1	236.8	2.3	2.3	1.9	0	8.2

Note: Mean, standard deviation (SD) and range of reading speed and accuracy for the 15 passages (A–O).

Abbreviation: wpm, words per minute.

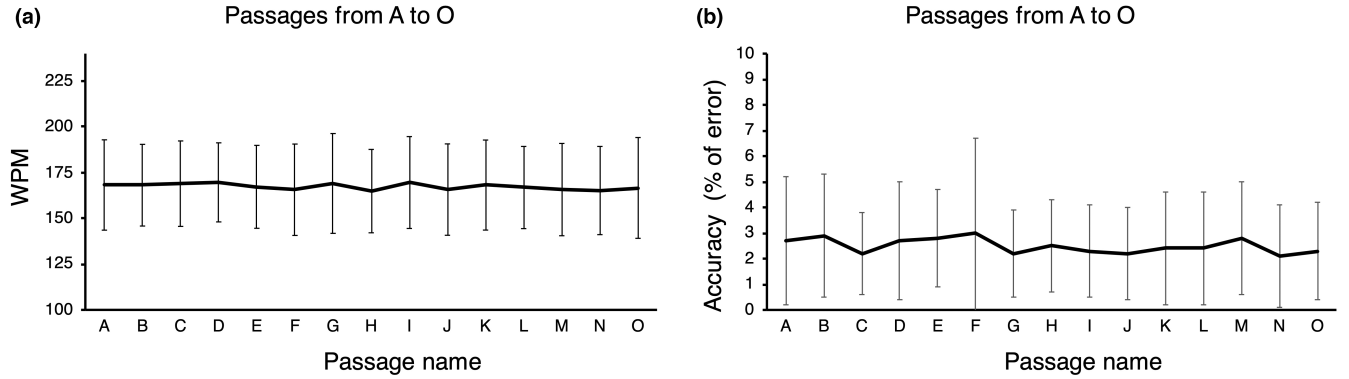

FIGURE 1 Equivalence between the passages. Mean (error bars, standard deviation) of (a) reading speed (words per minute, wpm) and (b) accuracy (% of error) for the 15 passages read by the complete sample ($N = 32$).

TABLE 3 Practice and fatigue effects.

	Reading speed (wpm)				Accuracy (percentage of errors)				
Reading order	**Mean**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**
First reading	155.2	24.2	113.3	223.3	3.9	2.7	4.2	0	19.8
Second reading	166.9	23.3	119.5	217.3	3.0	2.2	2.2	0	8.2
Third reading	166.3	23.9	125.0	220.7	2.6	2.7	1.5	0	6.8
Fourth reading	167.6	22.7	125.1	217.2	2.1	1.3	1.9	0	8.1
Fifth reading	169.0	24.3	114.0	217.8	2.3	1.5	1.8	0.3	7.3
Sixth reading	169.6	22.5	122.5	223.5	2.7	2.3	1.8	0	6.8
Seventh reading	170.1	23.5	118.2	223.1	2.4	2.0	1.7	0	6.8
Eighth reading	167.1	22.6	122.6	218.4	2.9	2.5	2.1	0	7.7
Ninth reading	168.1	25.1	116.0	232.5	2.3	1.8	2.1	0	8.5
Tenth reading	171.1	27.0	118.1	220.3	2.0	1.8	1.6	0	5.7
Eleventh reading	169.0	22.4	125.5	236.8	1.9	1.8	1.7	0	7.0
Twelfth reading	167.2	24.2	119.5	210.9	2.4	2.0	2.2	0	9.8
Thirteenth reading	166.3	25.0	115.5	215.2	2.1	1.3	2.1	0	7.7
Fourteenth reading	168.5	24.6	115.2	214.2	2.4	1.5	2.4	0	8.5
Fifteenth reading	167.7	25.4	120.5	221.2	2.6	2.3	1.7	0	8.2

Note: Mean, standard deviation (SD) and range of reading speed (words per minute, wpm) and accuracy for the readings, from the first to the fifteenth. Data are averaged separately by the presentation order of the passages.

Practice and fatigue effects

Descriptive statistics are reported in Table 3. Figure 2 shows reading speed and accuracy as a function of the reading order. Reading speed data based on reading order were normally distributed for all readings (Shapiro–Wilk test: $p > 0.05$) whereas reading accuracy was not normally distributed in 14 out of the 15 passages (Shapiro–Wilk test: all $p_s < 0.05$ except one, requiring non-parametric analysis). The repeated-measures ANOVA was significant for reading speed (F (1, 14) = 4.95; $p < 0.001$) indicating an effect of presentation order. Regarding reading accuracy, the Friedman ANOVA showed a significant difference between the readings ($r = 33.3$; $p = 0.003$).

Post-hoc testing (t-tests, two-tailed) for reading speed showed that the first reading was significantly slower than the subsequent 14 passages (all p values <0.001). All other paired comparisons (including those involving the last readings for the detection of fatigue effects) were not significant ($p > 0.05$), except for the 10th versus 13th and 10th versus 15th ($p < 0.05$). After Bonferroni correction for multiple comparisons (alpha lowered to 0.0005), the first reading showed a significantly slower rate than 12 out of the other 14 passages (the 12th and 13th did not differ significantly from the first after correction). The mean difference between the first passage presented and all subsequent ones was about 13 wpm. There were no other significant differences in reading speed between pairs. The mean difference between readings for all other presentations was approximately 1.7 wpm. The mean reading speed (excluding the first reading) was 168.2 wpm. Concerning reading accuracy, though the first reading was less accurate than

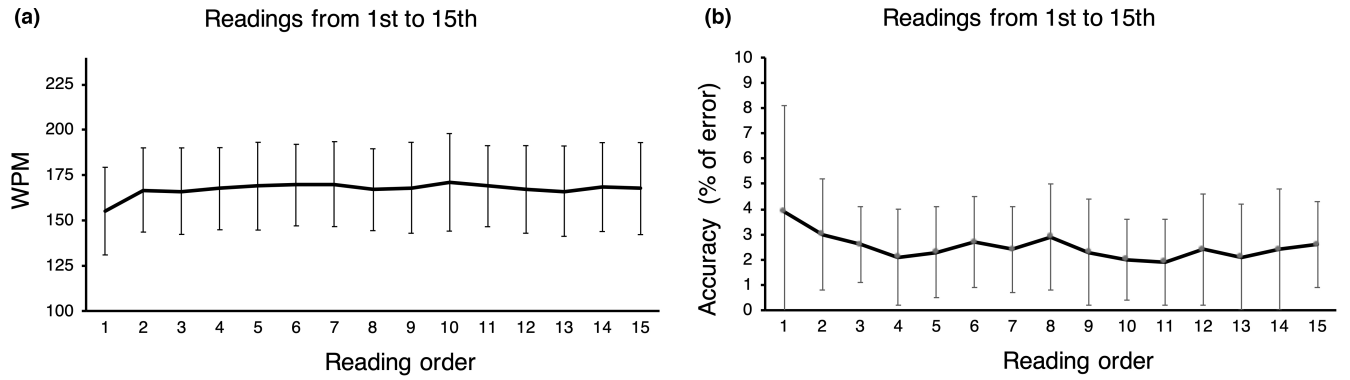

FIGURE 2 Practice and fatigue effects. (a) Reading speed (words per minute, wpm) and (b) accuracy (% of error) as a function of the reading order for the complete sample ($N = 32$).

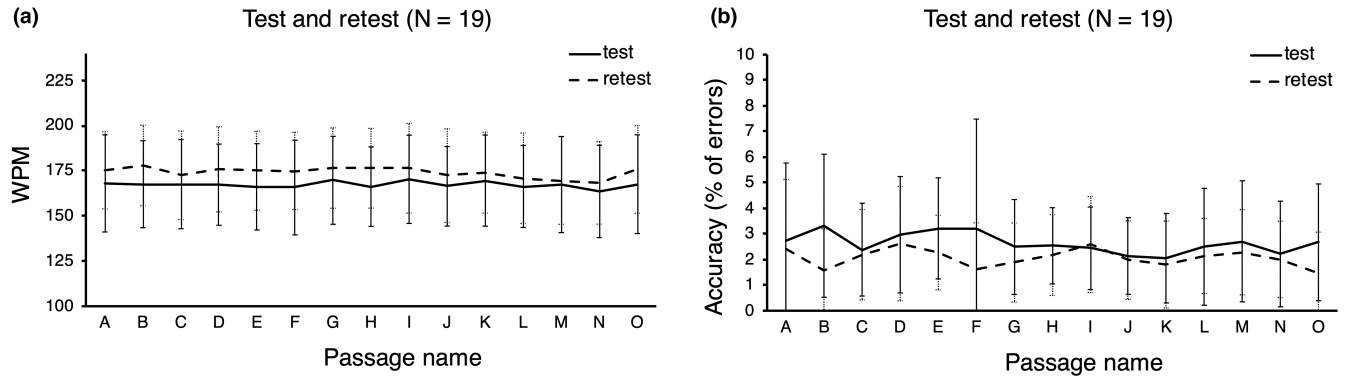

FIGURE 3 Equivalence between the passages: (a) Reading speed (words per minute, wpm) and (b) accuracy (% of errors) for the test and retest of the 15 passages of the Italian Wilkins Rate of Reading Test read by the sub-sample of 19 participants.

the following ones (see Table 3 and Figure 2), post-hoc testing with Dunn–Bonferroni correction for multiple comparisons did not detect any significant difference as a function of the reading order. The mean accuracy was more than 97%.

Test–retest reliability

Nineteen subjects (8 males and 11 females; mean age = 22.9 years, SD = 1.9, range = 20.1–26.5; mean years of formal education = 15.8 years, SD = 0.9, range 14–17 years) were retested 2 weeks after the first session. The results for reading speed and accuracy as a function of the 15 passages are shown in Figure 3 for test and retest as well as in Table 4, where the ICCs for reading speed and accuracy and statistical comparisons (paired tests) are provided.

The general trend in retest was a better performance for both reading speed and accuracy, as seen in 6 and 3 out of 15 for reading speed and accuracy, respectively; albeit with very few statistically significant differences. The ICC for reading speed (wpm), which is the usual test metric of the WRRT, was acceptable for all of the passages (0.67–0.90).

No Bland–Altman plots of the test–retest measurements of reading speed for the passages (A–O) showed any proportional bias (Figure 4 shows the Bland–Altman plot for the reading speed of passage C as an example). Table 5 shows the mean differences between the test and retest, the limits of agreement (LoAs) calculated as the mean difference ± 1.96 × SD of the difference and the CIs for the LoAs, as well as the correlation coefficients between the test–retest means. The absence of any proportional bias in the Bland–Altman plots is consistent with the correlation coefficients reported in Table 5, which were not significant for any of the passages. As for accuracy, Bland–Altman plots of the test–retest measurements did not show any proportional bias, consistent with non-significant correlations for 13 out of the 15 passages. The two significant correlations (passages B and O) indicate that for these two passages, the greater the percentage of errors (mean test–retest) the greater the percentage of errors at retest compared to the test (i.e., when the correlation was positive, greater accuracy values of the test–retest mean corresponded to a greater percentage of errors at the retest than the test).

DISCUSSION

The visual skills required for the dynamic task of reading involve many visual functions, including visual acuity,

TABLE 4 Test–retest ($N = 19$).

	Reading speed (words per minute)				Accuracy (percentage of errors)			
Passage name	Test Mean ± SD	Retest Mean ± SD	ICC (95% CI)	Comparison (*p*-value of paired *t*-test)	Test Mean ± SD	Retest Mean ± SD	ICC (95% CI)	Comparison (*p*-value of Wilcoxon)
A	168.0 ± 26.9	175.3 ± 21.4	0.67*** (0.32–0.86)	n.s	2.7 ± 3.0	2.4 ± 2.7	0.80*** (0.53–0.92)	n.s
B	167.6 ± 24.1	178.0 ± 22.3	0.75*** (0.45–0.89)	$p < 0.05$	3.3 ± 2.8	1.6 ± 1.6	0.40* (−0.06–0.72)	$p < 0.01$
C	167.6 ± 24.7	172.6 ± 24.5	0.90*** (0.77–0.96)	n.s	2.4 ± 1.8	2.2 ± 1.8	0.25 (−0.22–0.62)	n.s
D	167,3 ± 22.5	175.8 ± 23.6	0.87*** (0.69–0.95)	$p < 0.01$	3.0 ± 2.3	2.6 ± 2.2	0.47* (0.04–0.76)	n.s
E	166.1 ± 23.9	175.0 ± 21.8	0.83*** (0.61–0.93)	$p < 0.05$	3.2 ± 2.0	2.3 ± 1.5	0.53** (0.12–0.79)	$p < 0.05$
F	165.7 ± 26.2	175.0 ± 21.4	0.87*** (0.68–0.95)	$p < 0.01$	3.2 ± 4.3	1.6 ± 1.8	0.12 (−0.34–0.54)	n.s
G	169.7 ± 24.4	176.6 ± 22.3	0.77*** (0.50–0.91)	n.s	2.5 ± 1.9	1.9 ± 1.5	0.47* (0.03–0.76)	n.s
H	166.2 ± 22.0	176.4 ± 22.1	0.83*** (0.62–0.93)	$p < 0.01$	2.5 ± 1.5	2.2 ± 1.6	0.48* (0.05–0.76)	n.s
I	170.3 ± 24.4	176.4 ± 24.8	0.80*** (0.55–0.92)	n.s	2.4 ± 1.6	2.6 ± 1.9	0.20 (−0.27–0.59)	n.s
J	166.4 ± 22.0	172.4 ± 25.9	0.68*** (0.33–0.86)	n.s	2.1 ± 1.5	2.0 ± 1.5	0.27 (−0.19–0.64)	n.s
K	169.6 ± 25.2	174.0 ± 22.4	0.87*** (0.69–0.95)	n.s	2.1 ± 1.7	1.8 ± 1.7	0.08 (−0.37–0.51)	n.s
L	166.4 ± 22.7	170.9 ± 25.0	0.86*** (0.68–0.95)	n.s	2.5 ± 2.3	2.1 ± 1.5	0.57** (0.17–0.81)	n.s
M	167.4 ± 26.6	169.6 ± 24.3	0.82*** (0.58–0.93)	n.s	2.7 ± 2.4	2.3 ± 1.7	0.20 (−0.27–0.59)	n.s
N	163.6 ± 25.6	168.4 ± 22.9	0.84*** (0.63–0.94)	n.s	2.2 ± 2.1	2.0 ± 1.5	0.32 (−0.15–0.67)	n.s
O	167.6 ± 27.3	175.8 ± 24.3	0.86*** (0.67–0.94)	$p < 0.05$	2.7 ± 2.3	1.5 ± 1.6	0.64** (0.27–0.84)	$p < 0.01$

Note: Descriptive statistics of reading performance (speed and accuracy), Intraclass correlation coefficient (ICC) between test and retest measures for the 15 passages (A–O) calculated with two-way mixed effects model, consistency, single measures (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$), and *p*-values of paired comparison between test and retest.
Abbreviation: SD, standard deviation.

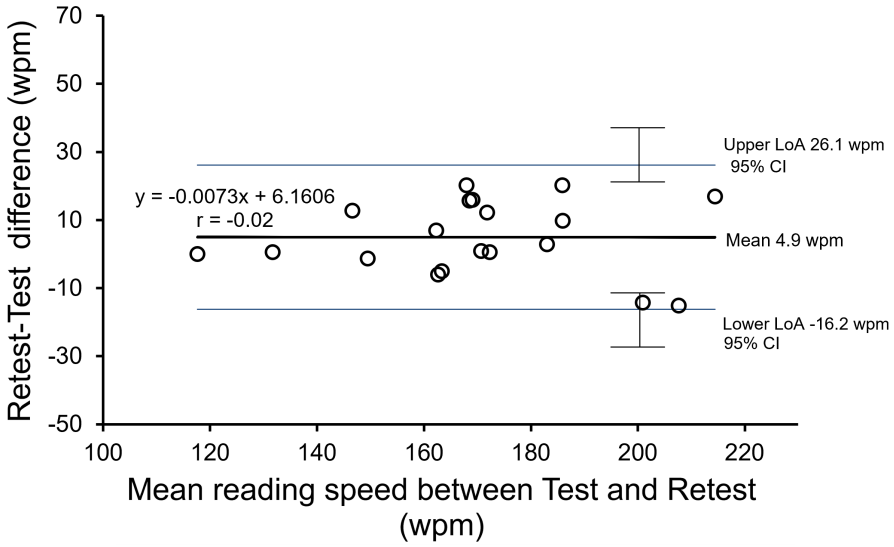

FIGURE 4 Example of Bland–Altman plot for the test–retest for one out of the 15 passages (passage C). The limits of agreement were calculated as $d_{mean} \pm 1.96\,SD_{diff}$. wpm, words per minute.

binocular co-ordination, ocular accommodation and eye movement control. The WRRT is designed to be sensitive to visual factors while requiring minimal language skills. The potential for several parallel versions of the test makes it particularly amenable to investigating the effect of optometric interventions. Therefore, it is not surprising that the test has been used in a variety of research topics, including visual stress, refractive corrections, head tilt, low-powered addition lenses, prismatic corrections, dry eye and low-dose atropine.[22–30]

As noted in the Introduction, when versions are created in different languages, these should be transliterations rather than translations. The Italian version of the WRRT created for this study, like the English version, uses simple commonly encountered words. This version contains more two-syllable words than the English WRRT, but fewer

TABLE 5 Test–retest findings ($N = 19$).

	Reading speed (wpm)		Accuracy (percentage of errors)	
Passage name	**Mean difference (LoA 95%) [CI for LoA: outer limits; inner limits]**	**Pearson correlation coefficient *r*; *p*-value**	**Mean difference (LoA 95%) [CI for LoA: outer limits; inner limits]**	**Spearman correlation coefficient *r*; *p*-value**
A	7.3 (−31.7–46.2) [−52.1–66.7; −22.7 – 37.3]	−0.30; n.s.	−0.3 (−3.9–3.3) [−7.1–6.5; −2.8 – 2.2]	−0.04; n.s.
B	10.4 (−22.1–42.8) [−38.9–59.7; −14.5 – 35.3]	−0.12; n.s.	−1.7 (−6.7–3.2) [−11.1–7.7; −5.2 – 1.8]	−0.55; $p < 0.05$
C	4.9 (−16.2–26.1) [−27.3–37.1; −11.4 – 21.2]	−0.02; n.s.	−0.2 (−4.5–4.1) [−8.4–8.1; −3.3 – 2.9]	−0.04; n.s.
D	8.5 (−14.6–31.6) [−26.7–43.7; −9.3 – 26.3]	0.09; n.s.	−0.3 (−4.9–4.2) [−9.0–8.4; −3.5 – 2.9]	−0.03; n.s.
E	8.9 (−17.4–35.2) [−31.1–48.9; −11.3 – 29.1]	−0.16; n.s.	−0.9 (−4.2–2.3) [−7.3–5.5; −3.3 – 1.5]	−0.17; n.s.
F	9.2 (−15.2–33.6) [−27.8–46.2; −9.5 – 27.9]	−0.40; n.s.	−1.6 (−10.1–7.0) [−18.2–15.0; −7.7 – 4.5]	−0.12; n.s.
G	6.8 (−24.0–37.6) [−40.1–53.7; −16.9 – 30.5]	−0.14; n.s.	−0.6 (−4.0–2.8) [−7.4–6.2; −3.1 – 1.9]	−0.24; n.s.
H	10.2 (−14.8–35.1) [−27.7–48.1; −9.0 – 29.7]	0.01; n.s.	−0.4 (−3.4–2.7) [−6.4–5.6; −2.6 – 1.8]	0.05; n.s.
I	6.1 (−24.6–36.8) [−40.8–53.0; −17.6 – 29.8]	0.03; n.s.	0.1 (−4.2–4.5) [−8.2–8.4; −3.0 – 3.2]	0.01; n.s.
J	5.9 (−32.0–43.9) [−52.0–63.8; −23.4 – 35.2]	0.22; n.s.	−0.2 (−3.7–3.4) [−7.0–6.6; −2.7 – 2.3]	0.20; n.s.
K	4.4 (−19.6–28.4) [−32.0–40.8; −14.0 – 22.8]	−0.24; n.s.	−0.2 (−4.8–4.3) [−8.9–8.5; −3.4 – 3.0]	−0.15; n.s.
L	4.5 (−20.0–29.0) [−32.8–41.8; −14.4 – 23.4]	0.19; n.s.	−0.4 (−3.8–3.1) [−7.2–6.4; −2.9 – 2.1]	−0.24; n.s.
M	2.2 (−28.0–32.5) [−43.8–48.2; −21.0 – 25.4]	−0.16; n.s.	−0.4 (−5.5–4.6) [−10.2–9.4; −4.0 – 3.2]	−0.23; n.s.
N	4.8 (−22.1–31.6) [−36.1–45.7; −15.9 – 25.5]	−0.20; n.s.	−0.2 (−4.3–3.9) [−8.1–7.7; −3.1 – 2.7]	−0.16; n.s.
O	8.2 (−18.9–35.3) [−33.0–49.4; −12.6 – 29.0]	−0.23; n.s.	−1.2 (−4.5–2.1) [−7.6–5.2; −3.6 – 1.2]	−0.48; $p < 0.05$

Note: Bland–Altman plot information for reading speed and accuracy for the 15 passages (A–O): mean difference between test and retest measures and limits of agreement (LoAs) calculated as mean difference ± 1.96 × standard deviation (SD) of the difference; confidence intervals (CI); calculated for LoA[36]; correlation coefficient between the mean of the test–retest measurements and the difference between these two measurements, also with the level of significance. A positive mean difference in reading speed indicates that the retest was faster than the test, whereas a negative mean difference in reading accuracy indicates the retest was more accurate than the test. Positive correlation coefficients for reading speed indicate that the quicker the reading speed (mean test–retest) then the quicker the retest compared to the test. Positive correlation coefficients for accuracy indicate the greater the percentage of errors (mean test–retest), then the greater the percentage of errors at retest compared to the test.

than the previous Italian version (6 vs. 13). This is partly due to the intrinsic characteristics of the Italian language (in which most four-letter high-frequency words are disyllabic). Another reason for this difference was the need to reduce the probability of grammatical, syntactic or semantic connections among words; this made some monosyllabic items unsuitable. In a test such as the WRRT, it is important to consider the visual layout once the simplicity of the words used has been guaranteed by choosing from high-frequency words. The main constraints to respect here were represented by critical parameters such as word length in terms of the number of characters, the total number of characters, font type, character size and word and line spacing. In the present Italian version, these parameters were all matched with the English version. Similarly, the International Reading Speed Texts (IreST),[31] a reading chart based on text paragraphs comprising equivalent versions for 17 languages, also shows a match for the number of characters but not syllables between the English and Italian versions. To summarise, the most important adherence to the English WRRT implies respect for the visuoperceptual features of the stimuli once the role of language has been minimised through very easy-to-read words.

It should be noted that reading speed in wpm varies both within a language (as a function of the study methodology) and across languages (as a function of both the methodology and language features). Several studies have indicated that the average adult English reading-aloud rate is around 185 wpm. For example, a meta-analysis by Brysbaert[32] reported an average reading rate of 183 wpm; Gilchrist and coworkers[6] reported 184 wpm in adult readers (mean age 21.8 years) with the WRRT. The literature also reveals that different reading rates characterise different cohorts of adult readers within the same study (e.g., 159 wpm in the second group of adult proficient readers in Gilchrist et al.[6]). The average reading speed (167.3 ± 1.6 wpm) achieved by our group of young university students (22.9 years, SD = 1.8, range 20.1–27.2) for the present Italian test lies within the range of the above-mentioned figures for English. Therefore, reading rate differences between the present Italian version and WRRT English studies are not surprising if one also considers the intrinsic features of the different languages (e.g., English and Italian are characterised by a deep and shallow orthography, respectively, and English syllables generally contain more letters than Italian). Similar differences in reading speed between Italian and English can also be found in the above-mentioned IreST[31] standardised reading chart. The paragraphs of this particular international chart were developed according to linguistic constraints to make the paragraphs equivalent across languages. Also, in the case of the IreST, a perfect linguistic match was not possible for the number of words and syllables, and word length, but it was matched for the number of characters per text. Interestingly, also in the IReST, the reading rate was slower for Italian (188 wpm) than for the English language (228 wpm). Of course, for both languages,

reading rates were faster for the IReST than the WRRT because the former is based on paragraphs with meaning instead of random words.

Since the WRRT involves successive testing with parallel versions of the test, it is important to investigate the effect of practice and fatigue on the repeatability of the test. Although the original version of the WRRT was typically carried out four times, twice with and twice without an intervention (or with a control condition), in some studies, parallel versions of the WRRT have been applied numerous times to the same observers.[33] Therefore, in the present research 15 versions were investigated. This study revealed a practice effect that became manifest only with the first passage read, where performance was slower by approximately 13 wpm on average. In contrast for reading accuracy, there were no significant differences as a function of the reading order. Therefore, because of the effects of reading order on reading speed, in case of repeated testing within the same session, it is important to allow practice with one passage before starting the performance measurements. Interestingly, a few previous studies with the WRRT have departed from the test instructions by having participants read part of one passage in a practice session. For example, Evans and Joseph[2] used two lines of the fourth passage for practice before testing started; other practice options included having participants read the first line from the fourth passage;[24,26] reading a passage for 30 s[25] or allowing participants a 'practice run'.[34] Other studies, typically those with younger children, have required subjects to read the 15 words used out loud, which would have provided some form of practice.[3,5] The practice effect with the first WRRT passage detected in the present research was first noted by Wilkins et al.[3] and acknowledged by other researchers.[10,35] It has been noted that since the first passage is typically read with the intervention, the practice effect will create a conservative bias against finding a beneficial effect of the intervention.[35] Nonetheless, the most recent (2019) version of the test includes a practice passage and the 2019 test instructions advises that the patient reads the practice passage first, before the formal test is commenced (ioosales.co.uk/). The present research supports this development.

Previous research has evaluated the test–retest repeatability of the WRRT.[3,6–9] The present study adds to this body of work by assessing the ICC and Bland–Altman plots in the Italian version. Koo and Li[20] described ICC estimates of values <0.50 as poor, between 0.50 and 0.75 as moderate, between 0.75 and 0.90 as good and >0.90 as excellent. Using this classification, all the estimates for the usual test metric (wpm) of the WRRT were good apart from passages A and J, for which the ICC estimation was graded as moderate (Table 4). Reading speed was significantly faster on retest for 6 of the 15 passages (Tables 4 and 5), but the mean difference only marginally exceeded a clinically significant difference of 10 wpm for two passages (i.e., B and H). Wilkins et al.[3] found a similar effect of practice with a quicker reading speed on retest of about 3%. Moreover, Bland–Altman plots did not show any proportional bias for reading speed in all test–retest comparisons, that is, the differences between test and retest measures were not affected by the reading speed (mean test–retest). Similarly, the clinical significance of the difference in reading accuracy, which resulted in a statistically significant improvement at the retest in three passages was negligible (between 1.2% and 1.6% of errors). Overall, the results indicate that the main indicator of performance for the RRT, namely reading speed measured as wpm, has good test–retest reliability along with a very low rate of errors.

To summarise, a novel feature of this study is the large number (15) of parallel versions of the Italian WRRT that were developed and evaluated. Some investigations have used a large number of repetitions of the test,[33] and the 15 parallel versions may be useful for other researchers. It is reassuring that the present work revealed no significant fatigue effects.

CONCLUSION

An Italian version of the WRRT produced highly repeatable results from a large set of passages. The present work confirmed a previous finding of a slight practice effect between the first and subsequent readings of the test, and as advised in the latest version of the test instructions, it is recommended that a practice attempt be allowed before the test is used.

AUTHOR CONTRIBUTIONS

Fabrizio Zeri: Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); resources (equal); software (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Silvia Tavazzi:** Conceptualization (equal); data curation (equal); funding acquisition (equal); project administration (equal); resources (equal); supervision (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Marco Punzi:** Data curation (equal); formal analysis (equal); investigation (equal); software (equal); writing – review and editing (equal). **Federica Miglio:** Data curation (equal); investigation (equal); writing –review and editing (equal). **Bruce J. W. Evans:** Writing – original draft (equal); writing – review and editing (equal). **Maria De Luca:** Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); software (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal).

ACKNOWLEDGEMENTS

The authors wish to thank Lara Cogato for helping in data collection. Open Access Funding provided by Universita degli Studi di Milano-Bicocca within the CRUI-CARE Agreement.

FUNDING INFORMATION

This work was partially supported by the Italian Ministry of Health (Ministero della Salute, Ricerca Corrente, Linea 1) to IRCCS Fondazione Santa Lucia.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest regarding the publication of this paper.

ORCID

Fabrizio Zeri https://orcid.org/0000-0003-0529-555X
Silvia Tavazzi https://orcid.org/0000-0001-5786-2021
Bruce J. W. Evans https://orcid.org/0000-0003-2824-4595
Maria De Luca https://orcid.org/0000-0003-0084-3772

REFERENCES

1. Wilkins A. Coloured overlays and their effects on reading speed: a review. *Ophthalmic Physiol Opt*. 2002;22:448–54.
2. Evans BJW, Joseph F. The effect of coloured filters on the rate of reading in an adult student population. *Ophthalmic Physiol Opt*. 2002;22:535–45.
3. Wilkins AJ, Jeanes RJ, Pumfrey PD, Laskier M. Rate of Reading test®: its reliability, and its validity in the assessment of the effects of coloured overlays. *Ophthalmic Physiol Opt*. 1996;16:491–7.
4. Northway N. Predicting the continued use of overlays in school children – a comparison of the Developmental Eye Movement test and the Rate of Reading test. *Ophthalmic Physiol Opt*. 2003;23:457–64.
5. Kriss I, Evans BJW. The relationship between dyslexia and Meares-Irlen syndrome. *J Res Read*. 2005;28:350–64.
6. Gilchrist JM, Allen PM, Monger L, Srinivasan K, Wilkins A. Precision, reliability and application of the Wilkins Rate of Reading test. *Ophthalmic Physiol Opt*. 2021;41:1198–208.
7. Wilkins A, Lewis E, Smith F, Rowland E, Tweedie W. Coloured overlays and their benefit for reading. *J Res Read*. 2001;24:41–64.
8. Jeanes R, Busby A, Martin J, Lewis E, Stevenson N, Pointon D, et al. Prolonged use of coloured overlays for classroom reading. *Br J Psychol*. 1997;88:531–48.
9. Hollingsworth RS, Ludlow AK, Wilkins AJ, Calver RI, Allen PM. Visual performance and the use of colored filters in children who are deaf. *Optom Vis Sci*. 2015;92:690–9.
10. Srinivasan K, Krishnan G, Wilkins A, Allen P. Reliability and validity of a Kannada rate of reading test. *Indian J Ophthalmol*. 2018;66:630–3.
11. Garcia ACO, Momensohn-Santos TM, Vilhena DDA. Effects of spectral overlays on reading performance of Brazilian elementary school children. *Folia Phoniatr Logop*. 2018;69:219–25.
12. Esmaeelpour M, Moest P, Krimpmann-Rehberg B. An evaluation of colored overlays and prismatic correction on the reading rate of elementary school children with scotopic sensitivity syndrome. *Optom Vis Sci*. 2002;79:102. E-abstract 17048.
13. Zeri F, Naroo SA, Zoccolotti P, De Luca M. Pattern of reading eye movements during monovision contact lens wear in presbyopes. *Sci Rep*. 2018;8:15574. https://doi.org/10.1038/s41598-018-33934-6
14. Zeri F, Berchicci M, Naroo SA, Pitzalis S, Di Russo F. Immediate cortical adaptation in visual and non-visual areas functions induced by monovision. *J Physiol*. 2018;596:253–66.
15. Virgili G, Cordaro C, Bigoni A, Crovato S, Cecchini P, Menchini U. Reading acuity in children: evaluation and reliability using MNREAD charts. *Invest Ophthalmol Vis Sci*. 2004;45:3349–54.
16. Calossi A, Boccardo L, Fossetti A, Radner W. Design of short Italian sentences to assess near vision performance. *J Optom*. 2014;7:203–9.
17. Zeri F, Beltramo I, Boccardo L, Palumbo P, Petitti V, Wolffsohn JS, et al. An Italian translation and validation of the near activity visual questionnaire (NAVQ). *Eur J Ophthalmol*. 2017;27:640–5.
18. De Mauro T. Nuovo vocabolario di base della lingua italiana [Internet]. 2016. Available from: https://www.internazionale.it/opinione/tullio-de-mauro/2016/12/23/il-nuovo-vocabolario-di-base-della-lingua-italiana. Accessed 15 May 2018.
19. Burns DH, Allen PM, Edgar DF, Evans BJW. Sources of error in clinical measurement of the amplitude of accommodation. *J Optom*. 2020;13:13–4.
20. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155–63.
21. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Int J Nurs Stud*. 2010;47:931–6.
22. Evans BJW, Allen PM. A systematic review of controlled trials on visual stress using intuitive overlays or the intuitive colorimeter. *J Optom*. 2016;9:205–18.
23. Firth AY, Machin J, Watkins CL. Tilt and reading speed. *J AAPOS*. 2007;11:52–4.
24. O'Leary CI, Evans BJW. Double-masked randomised placebo-controlled trial of the effect of prismatic corrections on rate of reading and the relationship with symptoms. *Ophthalmic Physiol Opt*. 2006;26:555–65.
25. Ludlow AK, Wilkins AJ, Heaton P. The effect of coloured overlays on reading ability in children with autism. *J Autism Dev Disord*. 2006;36:507–16.
26. O'Leary C, Evans B, Edgar D. The effect of low refractive corrections on rate of reading. *Optom Pract*. 2014;15:87–100.
27. Ridder WH, Zhang Y, Huang JF. Evaluation of reading speed and contrast sensitivity in dry eye disease. *Optom Vis Sci*. 2013;90:37–44.
28. Ousler GW, Rodriguez JD, Smith LM, Lane KJ, Heckley C, Angjeli E, et al. Optimizing reading tests for dry eye disease. *Cornea*. 2015;34:917–21.
29. Loughman J, Flitcroft D. The acceptability and visual impact of 0.01% atropine in a Caucasian population. *Br J Ophthalmol*. 2016;100:1525–9.
30. Yammouni R, Evans BJW. Is reading rate in digital eyestrain influenced by binocular and accommodative anomalies? *J Optom*. 2021;14:229–39.
31. Trauzettel-Klosinski S, Dietz K. Standardized assessment of reading performance: the new International Reading Speed Texts IReST. *Invest Ophthalmol Vis Sci*. 2012;53:5452–61.
32. Brysbaert M. How many words do we read per minute? A review and meta-analysis of reading rate. *J Mem Lang*. 2019;109:104047. https://doi.org/10.1016/j.jml.2019.104047
33. Wilkins A, Sihra N, Smith IN. How precise do precision tints have to be and how many are necessary? *Ophthalmic Physiol Opt*. 2005;25:269–76.
34. Yammouni R, Evans BJ. An investigation of low power convex lenses (adds) for eyestrain in the digital age (CLEDA). *J Optom*. 2020;13:198–209.
35. Allen PM, Dedi S, Kumar D, Patel T, Aloo M, Wilkins AJ. Accommodation, pattern glare, and coloured overlays. *Perception*. 2012;41:1458–67.
36. Carkeet A. A review of the use of confidence intervals for Bland-Altman limits of agreement in optometry and vision science. *Optom Vis Sci*. 2020;97:3–8.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zeri F, Tavazzi S, Punzi M, Miglio F, Evans BJW, De Luca M. New Italian version of the Wilkins Rate of Reading Test: Materials for repeated-measure designs in optometry and neuropsychological research. *Ophthalmic Physiol Opt*. 2023;00:1–11. https://doi.org/10.1111/opo.13134