African Vision and Eye Health
ISSN: (Online) 2410-1516, (Print) 2413-3183

Design and reliability of an optometric isiZulu paediatric rate of reading test: A case study

Authors:
Urvashni Nirghin[1]
Rekha Hansraj[1]
Lavanithum Joseph[2]

Affiliations:
[1]Department of Optometry, Faculty of Health Sciences, University of KwaZulu-Natal, Durban, South Africa

[2]Department of Audiology, Faculty of Health Sciences University of KwaZulu-Natal, Durban, South Africa

Corresponding author:
Urvashni Nirghin,
nirghinu@ukzn.ac.za

Dates:
Received: 21 May 2025
Accepted: 21 Oct. 2025
Published: 14 Jan. 2026

How to cite this article:
Nirghin U, Hansraj R, Joseph L. Design and reliability of an optometric isiZulu paediatric rate of reading test: A case study. Afr Vision Eye Health. 2026;85(1), a1077. https://doi.org/10.4102/aveh.v85i1.1077

Background: IsiZulu is an official language in South Africa and the primary home language (HL) for many South African children. Currently, there are no optometric reading tests specific for isiZulu-speaking children.

Aim: This study aimed to develop and evaluate an isiZulu Paediatric Rate of Reading (PRR) Test for children, with or without visual impairment.

Setting: The study was conducted in primary schools of KwaZulu-Natal, South Africa, focusing on isiZulu-speaking primary school children.

Methods: The study followed a mixed-methods design. High-frequency isiZulu words from primary school textbooks were used to create six versions of the PRR Test, matching visual acuities from 1.0M to 4.0M. Reliability was tested on 162 isiZulu-speaking primary school children, conveniently selected from HL and First Additional Language (FAL) schools, whereby two examiners recorded results for reading performance over two sessions, one week apart.

Results: Bland-Altman plots indicated strong inter-rater agreement with paired *t*-tests showing no significant differences ($P > 0.05$) in inter-rater reading parameters. However, statistical differences ($P < 0.05$) for inter-session agreement were observed for the majority. Intraclass correlations for reading measures ranged from moderate to excellent (0.848–1.000).

Conclusion: The test showed good reliability for reading measures, except that reading error variation was observed.

Contribution: This study contributes to the development of an isiZulu PRR Test, which is a reliable and linguistically relevant tool for assessing reading abilities in isiZulu-speaking children.

Keywords: reading; reading rate; speed; time; errors; visual impairment; children; optometry.

Introduction

Reading assessment has been a key focus in paediatric optometry examination and clinical research, aiming to identify and manage visual factors contributing to reading problems. The availability of optometric reading tests in a range of languages caters to the needs of global diversity. Apart from the accessibility of various English-based optometric tests,[1,2,3,4,5] to ensure inclusivity, other language-based reading tests have been designed. Several tests have subsequently been created, including but not limited to the following: MNREAD Chart, RADNER Reading Charts in Spanish, French, German, Dutch, Swedish, Danish, Hungarian and Italian,[6] Arabic continuous text reading chart,[7] Chinese reading chart for children,[8] Kannada rate of reading test,[9] and recently the Italian version of the Wilkins Rate of Reading Test.[10] However, most of these tests are not specifically designed for children, particularly primary school children, with the words being more applicable to the adult population.

It is for this reason that the optometric English Paediatric Rate of Reading (PRR) Test was designed by Nirghin for English-speaking primary school children in South Africa.[11] The test was developed primarily for vision-related assessments, particularly to evaluate the effects of visual and oculomotor deficits on reading rate, and has been shown to be both reliable and valid. It was designed using similar principles to those adopted by the Wilkins Reading Rate Test,[4] but with word selections drawn from prescribed textbooks used by South African primary school children.[11] Furthermore, the design incorporated a range of visual acuity levels to accommodate children with or without visual impairment. Given its relevance and proven reliability, the English PRR Test forms the basis for the development of the isiZulu version used in this study.

There are 12 official languages in South Africa (SA), with isiZulu being spoken by 22.7% of the population, and it is the most common home language (HL) spoken by children in SA.[12,13,14,15] It would be beneficial to have suitable assessments for evaluating the functional performance of isiZulu-speaking children in order to address the specific requirements of this demographic in a clinical context. Considering the recommendation of a prior study, the direct translation of previous tests should not be adopted, as it does not ensure linguistic frequency, word length and structural constraints necessary for test validity.[10] As this can compromise the accuracy of the test, reduce cross-language comparability and lead to misinterpretation of reading ability because of linguistic bias, this study aimed to design a rate of reading test specifically for isiZulu primary school children, whether isiZulu is their HL or first additional language (FAL), and to assess its reliability and validity. Similar to the English PRR Test,[11] different versions were to be considered for individuals with normal vision as well as those exhibiting visual impairment. For this publication, only the design and reliability of the optometric isiZulu PRR Test will be presented, with validity to be addressed in a subsequent publication.

The availability of reading tests for the assessment of reading rate, time or speed and accuracy considering the errors made, will be useful in a clinical context as well as exploring further research in the African continent where such research is very limited and possibly restricted by the language barrier.

Research methods and design

This study was approved by the University of KwaZulu-Natal's Biomedical Research Ethics Committee (BREC) and the Department of Basic Education, with all stipulated guidelines and the tenets of the Declaration of Helsinki adhered to. Informed consent and assent were obtained from the parents and participants, respectively. Data are available upon reasonable request and is subject to compliance with South African legal requirements, including the *Protection of Personal Information Act* (*POPIA*).

The study used a combination of qualitative and quantitative methods in its design. The qualitative aspect focused on creating the tool and analysing word frequency from prescribed readers of HL and FAL isiZulu-speaking learners. The reliability in this case study was assessed using a quantitative, observational-descriptive approach by evaluating reading rate across two examiners (inter-rater) over two sessions (inter-session) and one week apart.

Data collection

Development of the optometric isiZulu paediatric rate of reading test

The design of the isiZulu PRR Test closely follows the structure and procedures of the English PRR Test developed by the first author,[11] with the primary difference being the use of words sourced from isiZulu textbooks commonly used in South African primary schools. The test also shares similarities with the Wilkins Reading Rate Test,[4] but differs in several key aspects such as: the number of words included in each paragraph, the choice of words in terms of language and the inclusion of versions to accommodate different levels of visual acuity. While the overall test format remains consistent with the English version, these adaptations ensure that the isiZulu PRR Test is linguistically and visually appropriate for isiZulu-speaking learners. Considering the intent of the optometric isiZulu PRR Test, the selection of the words, their placement in each version of the test, visual acuity letter size and print scaling, font type, contrast and the design of the pre-test chart and recording scoresheet will be discussed.

Grade one isiZulu prescribed textbooks in Portable Document Format (PDF) were imported into the NVIVO qualitative data analysis software (version 11), and by using the Word Frequency Query in the program, a list of frequently used words was generated. Words with only three to four letters were selected to ensure an average level of word difficulty, as in the case of the English PRR Test.[11] Numbers, English words, punctuations, special characters, abbreviations and common nouns were excluded. From the generated list of 12 words, 10 words were randomly chosen, arranged left to right, consistent with the reading direction of isiZulu and assigned to 10 rows per paragraph. A total of six paragraphs were produced with each paragraph corresponding to one of six test versions (A–F). These paragraphs were positioned centrally on each version, without any outlining borders. Although no physical borders were drawn, consistent white space was maintained around the text, equal in thickness on the top, bottom, left and right to standardise the visual presentation across all test versions. Furthermore, consistent spacing between texts and lines in all versions was maintained using computer-generated print to avoid the impact of the crowding effect on reading as suggested by previous studies.[16,17]

The placement of words within each paragraph was determined using a randomisation method: each of the 12 words was assigned a number and a list of random numbers was used to guide 10-word placements per line. This ensured that each line and paragraph was unique in its word order. Consideration of the foregoing in the design was to mimic eye movement during reading, standardise reading difficulty, minimise the influence of linguistic factors and standardise the length and level of difficulty per line and version.

Each of the six versions (A–F) of the chart included text to measure near visual acuities of 1.0M, 1.2M, 1.6M, 2M, 3.2M and 4M, respectively, with notations in Snellen and Logarithm of Minimum Angle of Resolution (LogMAR) also included. Print smaller than 1.0M is not commonly found in books of primary school children, hence it was not included in the design, while relative distance magnification could be applied if print larger than 4M is required.[11] To ensure the appropriate scaling of the print for each test version, a graphic designer was enlisted and the formula $x = d\ tanq$ was applied, where x is the letter height (in mm), d is the viewing distance (in mm) and θ is the angle (5 minutes of arc) that each letter subtends at the eye.[18] The print size was maintained uniformly within each paragraph or test version and did not

vary between lines within a paragraph. The scaling excluded the ascenders and descenders of letters such as b, d, g and h. Du Toit[17] recommended that the typeface chosen for children's reading tests should reflect what they are commonly exposed to, hence the Quick Sand font typeface was considered to replicate prescribed textbooks and the teaching of writing skills of South African primary school children. Versions A to E were printed on individual 29.7 centimetre by 19.5 cm (A4) matte cardboard, while Version F was on 36.0 cm by 19.5 cm cardboard, considering the large print size. Complying with the contrast specifications in research, all text was printed in black on a white background.[19] For each version of the test (A–F), corresponding scoresheets were designed to provide the examiner a simultaneous opportunity to follow the reader while documenting errors and the time necessary for calculating the reading rate.

A pre-test chart was designed to ensure the child is familiar with the words before administering the test irrespective of the version used. The selected 12 words were presented on a matte white A4 cardboard, organised in two columns comprising six rows each. The words were printed in large font (5.82 mm) corresponding to an acuity level of 4.0M, in black ink with a white background to ensure legibility for children with a range of visual acuity levels.

Reliability of the optometric isiZulu paediatric rate of reading test

Before implementing a newly developed instrument, it is essential to establish the tool's reliability.[20] The process of assessing the reliability (inter-rater and inter-session) of the optometric isiZulu PRR Test is explained in this section.

Participants and sampling

From a list of public mainstream primary schools, with isiZulu being offered as the HL and FAL, two schools were selected by simple random sampling. Learners from Grades one to seven (ideally corresponding to ages 6–12 years) were recruited, as the isiZulu PRR Test was specifically designed for this primary school age range. The total number of normally-sighted participants in the study was determined to be 162 with 102 from isiZulu HL and 60 from FAL schools and was based on an expected power of 0.8, an effect size of 0.5 and a significance level of 0.05,[21] with the participants being selected through convenience sampling. The first 162 isiZulu-speaking children who agreed and received consent to participate in the study, passed the pre-test, pre-visual screening and had no systemic illness or special needs according to the school records, were included.

Study setting

The study was conducted in a three-metre by four-metre classroom, with consistent lighting provided by overhead fluorescent lights and natural light entering through windows located away from the testing area. Although the exact lux level was not measured, all fluorescent fixtures were operational to maintain uniform illumination throughout data collection. Only the two researchers and one participant were present in the room at any point in time, aiming to minimise distractions and create a comfortable environment for the participant to read without external pressures.

Accuracy and consistency of results

The researchers, one English-speaking and the other isiZulu-speaking, followed all protocols within the scope of optometric practice. To enhance standardisation, a detailed step-by-step guide was accessible to the researchers throughout data collection, and each examiner focused on a specific aspect of the research to ensure reliable and valid results.

Preliminary procedures

The learners underwent the optometric isiZulu PRR Test pre-test visual screening and pre-screening as the first steps and those who passed both these initial procedures proceeded to the reading rate evaluation for the reliability of the optometric isiZulu PRR Test.

- *Pre-test visual screening:* Initial screening aimed to eliminate visual problems that could impact reading rate results and included near visual acuity with or without habitual correction, near point of convergence (NPC) using a penlight and red and/or green spectacles cover test, accommodative response, amplitude of accommodation (push-up test), ocular-motility and internal ocular health assessments for all learners. Near visual acuity was assessed using a near logMAR chart at 40 cm, with monocular and binocular acuities recorded in logMAR to ensure that the print size used in the optometric isiZulu PRR test was appropriate. Learners who met all screening criteria proceeded to the reading pre-screening, while those who did not were referred to an eye care practitioner for further assessment. The screening criteria included near visual acuity of 1.0M (0.4 logMAR) or better, both monocularly and binocularly; NPC break and recovery values of 2.5 cm ± 4 cm and 4.5 cm ± 5 cm respectively; phoria within normal limits (1Δ exophoria ± 2Δ at distance and 3Δ exophoria ± 3Δ at near); an accommodative response of +0.50 D ± 0.25 D; amplitude of accommodation consistent with Hofstetter's formula for age; ocular-motility within Northeastern State University College of Optometry (NSUCO) age- and sex-based norms for ability, accuracy and head and/or body movement; and the absence of ocular pathology.[22]
- *isiZulu PRR Test pre-screening:* This screening aimed to identify learners capable of reading the isiZulu words in the test. Seated comfortably at a desk, the learner was presented with the pre-test chart, which was positioned 40 cm from their eyes and maintained using a 40 cm string. As the learners were higher than the desk, their reading posture naturally resulted in a downward gaze, typical of normal reading. The chart was viewed binocularly either unaided or with habitual correction, as each learner's visual acuity was appropriate for the acuity level of the pre-test chart (4.0M). Learners who successfully read all words proceeded to the reading rate evaluation to establish the reliability of the test.

Procedure for establishing the reliability of the optometric isiZulu paediatric rate of reading test

The necessary testing materials included the isiZulu PRR Test and scoresheet, a stopwatch, a 40 cm string and pens which were available throughout the data collection. Test instructions to each participant were provided in isiZulu and English, and participants were advised not to use their fingers to track words while reading. To ensure consistency with the learners' schooling materials, version D (2.0M) of the isiZulu PRR test was considered in measuring the reliability of the isiZulu PRR Test. Furthermore, none of the participants had visual acuities incompatible with the chosen version, which was less than 2.0M.

The participant was instructed to read the test chart placed on the desk positioned 40 cm away with the distance maintained with a 40 cm string connected to the chart. Participants were directed to view the chart with both eyes open, adopting a natural downward reading gaze and to read the words aloud as quickly and accurately as possible. Upon being signalled to begin, the participants commenced reading, and the stopwatch was started. Upon finishing the 100 words, the examiners ended the timing. The examiners were able to follow along with the participant using the corresponding version's scoresheet, with each examiner recording on an individual scoresheet. Errors, including additions, omissions, transpositions or substitutions were noted.[11,23] Based on the following formula (Equation 1),[4,11,24] reading rate was determined and recorded for each participant on their respective scoresheet.

$$\text{Reading Rate} = \frac{60 \text{ x Total number of words read correctly}}{\text{Total time taken in seconds}} \quad \text{[Eqn 1]}$$

The second reading rate evaluation on the same 162 participants was determined one week later (re-test or session two) using the same testing procedure as discussed. The result for each participant was documented on the same scoresheet per examiner but with a different coloured ink pen to identify test and re-test recordings. A one-week interval between the test and re-test, consistent with previous studies,[11,25] was used to minimise potential improvements in reading rate because of learning effects from prior exposure to the text. While improvements in reading performance may occur over time, this interval is considered sufficissent to mitigate immediate learning effects that could artificially enhance reliability estimates.

Data management and analysis

The demographic data of the participants and the test and re-test findings per participant per examiner were recorded directly from the scoresheet onto the Statistical Package for Social Sciences (SPSS) (version 23) program (SPSS Inc., Chicago, Illinois, US). The inter-rater and inter-session analysis was based on the Bland and Altman method, one-sample *t*-test ($P < 0.05$), and the Intraclass Correlation Coefficient (ICC) at a 95%confidence interval (CI). The classification of ICC levels was as follows: (1) poor reliability < 0.5, (2) moderate reliability 0.50–0.75, (3) good reliability 0.75–0.90, and (4) excellent reliability > 0.90.[20,26]

Ethical considerations

Ethical approval to conduct this study was obtained from the University of KwaZulu-Natal Biomedical Research Ethics Committee (No. BFC207/16).

Results

The reliability of the optometric isiZulu PRR Test among normally-sighted learners was assessed using Version D of the test through inter-rater and inter-session measurements. Participants' ages ranged from 6 years and 4 months (76 months) to 14 years and 8 months (176 months), with a mean age of 10 years and 3 months (123 months). The majority of participants were female (68%) and the Grades spanned from two (6.8%) to seven (25.9%). There were no participants from Grades two and four at the isiZulu FAL school, and Grade one from both schools as none of the learners returned the consent forms (Table 1).

Inter-rater reliability at the isiZulu home language and first additional language schools for the Z and E examiners

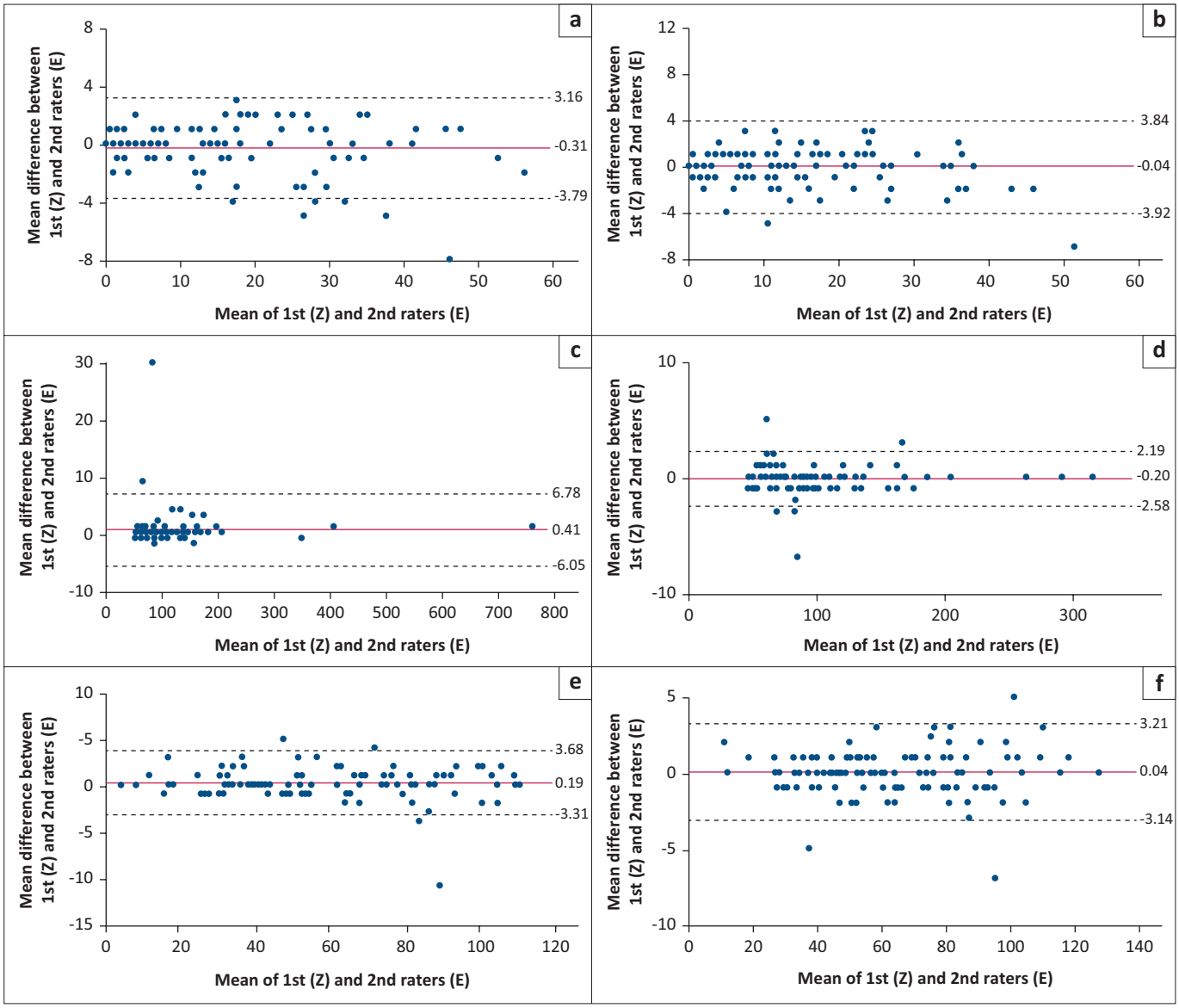
The inter-rater reliability measured the agreement between two examiners, one isiZulu-speaking (Z) and the other English-speaking (E) on the same participants at the isiZulu HL and FAL schools using version D of the optometric isiZulu PRR Test. Inter-rater reliability for measurements of total errors, time, and reading rate was assessed on day one (session one – Z1:E1) and again on day seven (session two – Z2:E2). The results were analysed using the Bland and Altman method (Figure 1 and Figure 2) in which the three horizontal lines represent the mean (red solid line) and the upper and lower Limits of Agreement (LoA) with mean ± 1.96 standard deviation (SD) respectively, are represented by dashed lines and set at a 95% CI.

The Bland and Altman plots depicted in Figure 1 and Figure 2 show that the majority of the categories of total errors, time, and reading rate for both sessions (Z1:E1 and Z2:E2) at the isiZulu HL and FAL schools fall on and within the 95%LoA. The exceptions are for total errors in session one at the HL school (94%, 96/102) and session two at the FAL school (93%, 56/60). These results demonstrate good inter-rater reliability

TABLE 1: The frequency distribution of participants by grade and school.

Grade	School					
	FAL		HL		Valid total	Cumulative
	n	(%)	*n*	(%)	*n*	%
2	0	0.00	11	10.78	11	6.80
3	20	33.33	20	19.61	40	24.70
4	0	0.00	21	20.59	21	13.00
5	20	33.33	20	19.61	38	23.50
6	0	0.00	10	9.80	10	6.20
7	20	33.33	20	19.61	42	25.90
Valid total	**60**	**100.00**	**102**	**100.00**	**162**	**100.00**

FAL, first additional language; HL, home language.



Note: (a) Total errors (Z1:E1) - version D; (b) total errors (Z2:E2) - version D; (c) total time (s) (Z1:E1) - version D; (d) total time (s) (Z2:E2) - version D; (e) reading rate in words per minute (wpm) (Z1:E1) - version D; (f) reading rate (wpm) (Z2:E2) - version D.

FIGURE 1: Inter-rater agreement represented by the Bland and Altman method with mean difference for total errors, total time, and reading rate plotted against their means one week apart (Z1:E1 and Z2:E2) between two examiners (E and Z) for the isiZulu PRR Test on HL school children.

for all measures, particularly total time and reading rate, for both sessions (Z1:E1 and Z2:E2), regardless of whether it is an HL or FAL school.

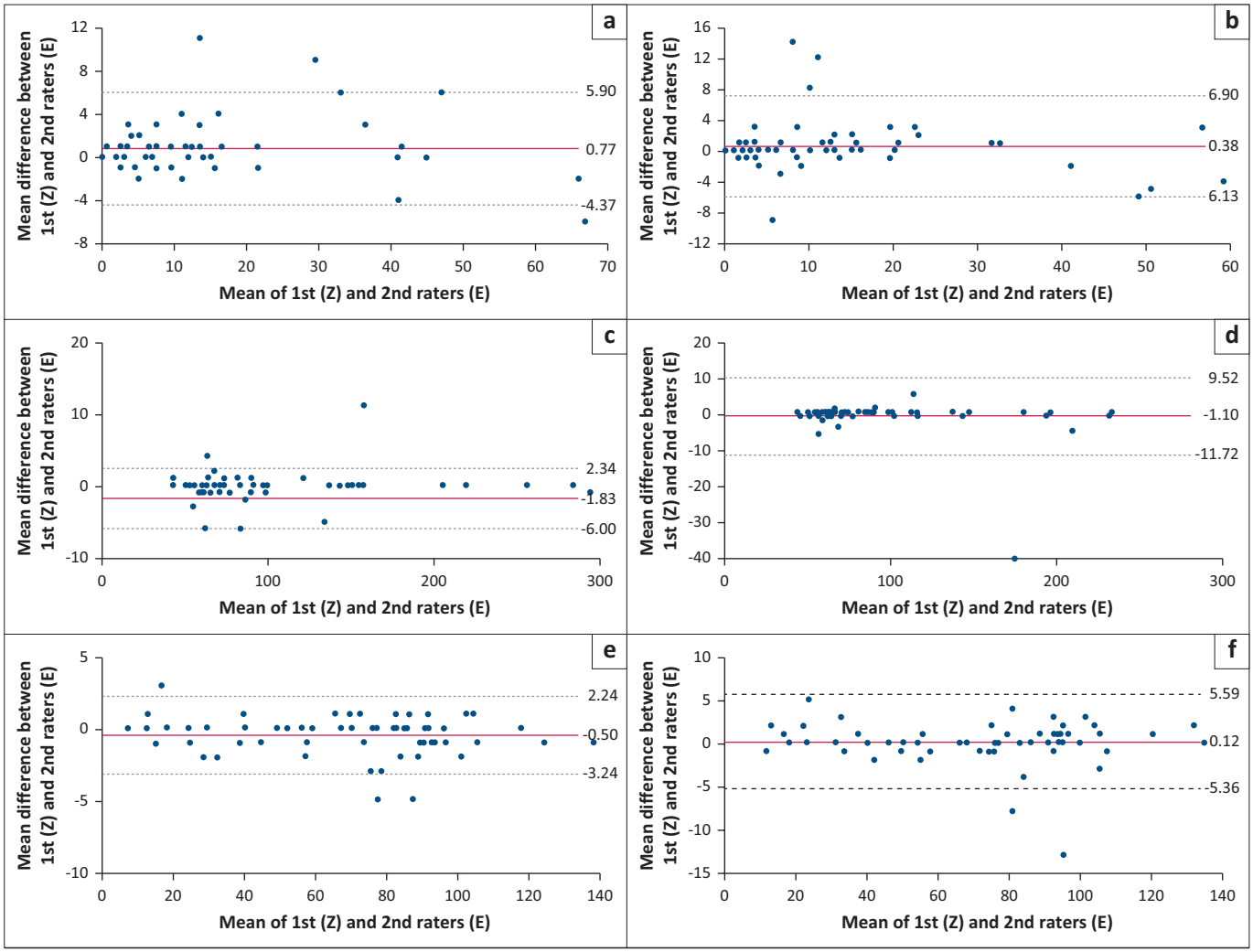
The ICC and one-sample *t*-test were performed to assess inter-rater reliability for total errors, time and reading rate by two examiners (referred to as Z and E) during session one (Z1:E1) and session two (Z2:E2) at HL and FAL schools. The results are presented in Table 2.

As shown in Table 2, the ICCs for all variables were ≥ 0.986, indicating minimal difference between results for total errors, time, and reading rate between the Z and E in both the HL and FAL schools, except for total errors on session one at the FAL school (0.848). This demonstrates an excellent linear relationship between the Z and E examiners for total errors, time and reading rate across both schools and sessions, with a good relationship for total errors in session one. The *P*-values

TABLE 2: The intraclass correlation and one-sample *t*-test for inter-rater reliability of the isiZulu Paediatric Rate of Reading Test on isiZulu home language and first additional language school children.

School	Session	Reading performance	Intraclass correlation coefficient			One-sample *t*-test	
			ICC	Lower bound	Upper bound	Mean difference	*P*-value
HL		Total errors	0.992	0.988	0.995	-0.31	0.077
	1	Total time (s)	0.999	0.999	1.00	0.41	0.204
		Reading rate (wpm)	0.998	0.997	0.999	0.04	0.807
		Total errors	0.986	0.991	0.998	-0.04	0.842
	2	Total time (s)	1.000	1.000	1.000	-0.19	0.107
		Reading rate (wpm)	0.998	0.997	0.999	0.20	0.294
FAL		Total errors	0.848	0.756	0.907	0.77	0.027
	1	Total time (s)	0.999	0.999	1.000	-1.83	0.507
		Reading rate (wpm)	0.999	0.998	0.999	-0.50	0.007
		Total errors	0.975	0.958	0.985	0.38	0.375
	2	Total time (s)	0.993	0.989	0.996	-1.10	0.121
		Reading rate (wpm)	0.996	0.993	0.998	0.12	0.747

ICC, intraclass correlation coefficient; FAL, first additional language; HL, home language; wpm, words-per-minute.



Note: (a) Total errors (Z1:E1) - version D; (b) total errors (Z2:E2) - version D; (c) total time (s) (Z1:E1) - version D; (d) total time (s) (Z2:E2) - version D; (e) reading rate in words per minute (wpm) (Z1:E1) - version D; (f) reading rate (wpm) (Z2:E2) - version D.

FIGURE 2: Inter-rater agreement represented by the Bland and Altman method with mean difference for total errors, total time and reading rate plotted against their means one week apart (Z1:E1 and Z2:E2) between two examiners (E and Z) for the isiZulu paediatric rate of reading test on first additional language school children.

for all reading measures ($P > 0.05$) indicate no significant difference in the means between the isiZulu-speaking examiner (Z) and English-speaking examiner (E) in both HL and FAL schools, except for total errors ($P = 0.027$) and reading rate ($P = 0.007$) for session one in the FAL school. There is a reduction in the mean difference for total errors and total time between both examiners (Z and E) in session two (Z2:E2) at both schools as illustrated in Figure 1 and Figure 2, except for reading rate. Examiner E made more errors in the HL school, while examiner Z recorded more errors in the FAL school.

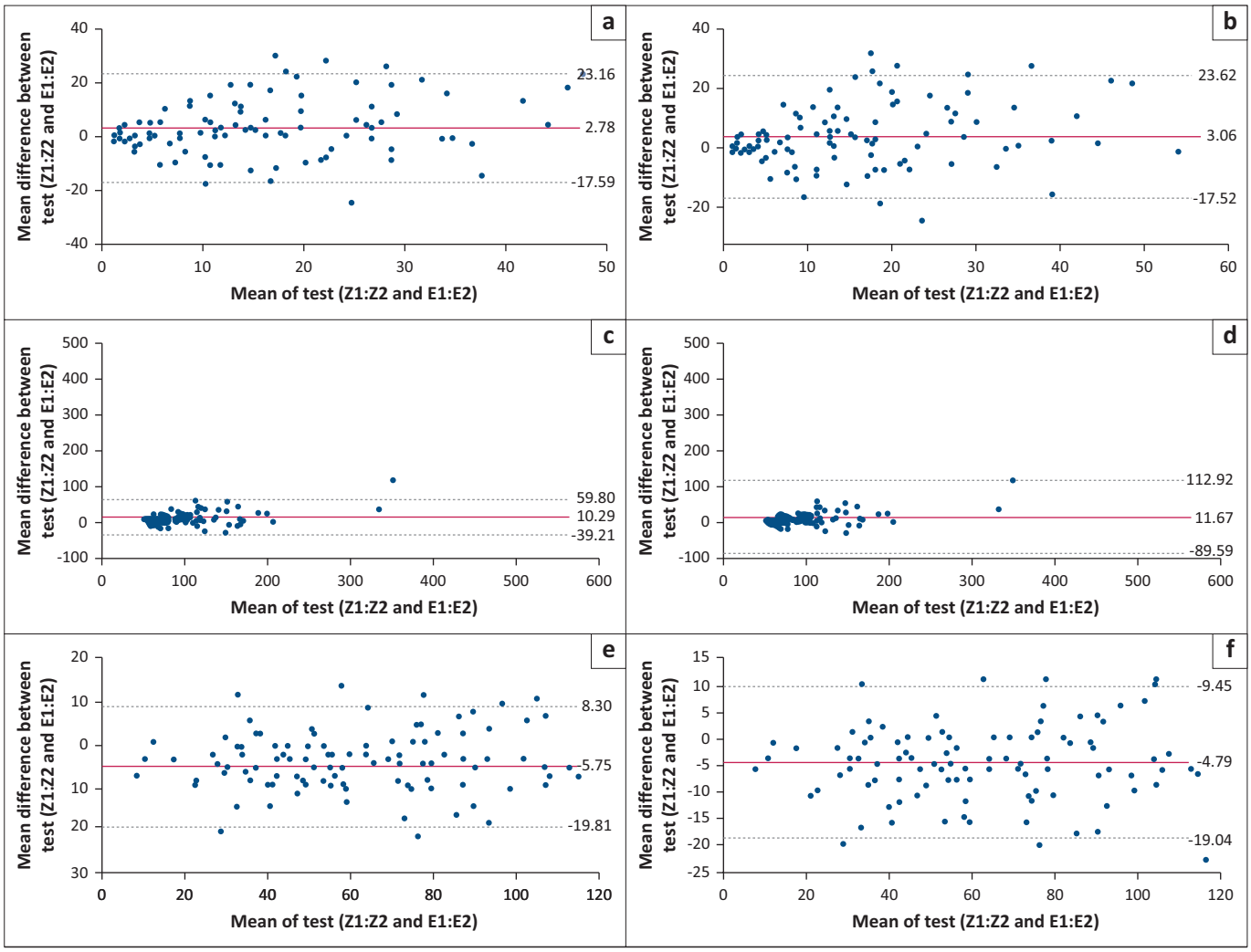
In addition, the upper and lower 95% LoA for all reading measures show an expansion in the range for errors, but reduced ranges for total time and reading rate in session two in the HL school.

Inter-session reliability at the isiZulu home language and first additional language schools for the Z and E examiners

In this study, inter-session reliability pertains to the consistency of test results over two sessions conducted by the same examiner on the same participants, separated by a specific period.[27] The analysis focused on total errors, time, and reading rate measured by examiners Z and E during two sessions: Z1 and E1 on day one, and Z2 and E2 one week later, using version D of the isiZulu PRR Test. The data were evaluated using the Bland and Altman method (see Figure 3 and Figure 4), ICC, and one-sample *t*-test (refer to Table 3) at a 95% CI.

The Bland and Altman plots (Figure 3 and Figure 4) illustrate that the inter-session results for examiners Z (Z1:Z2) and E (E1:E2) lie within the 95% LoA as follows; for total errors, 94%(96/102) and 95% (97/102) in the HL school, and 93% (56/60) and 95% (57/60) in the FAL; for time 98% (100/102) and 99% (101/102) in the HL school, and 93% (56/60) and 92% (55/60) in the FAL school; and for reading rate 93% (95/102) and 91% (93/102) in the HL school, and 95% (57/60) and 97% (58/60) in the FAL school, respectively.

The ICC and one-sample *t*-test findings for all reading measures for examiners Z and E for both sessions (Z1:Z2



Note: (a) Total errors (Z1:Z2) - version D; (b) total errors (E1:E2) - version D; (c) total time (s) (Z1:Z2) - version D; (d) total time (s) (E1:E2) - version D; (e) reading rate in words per minute (wpm) (Z1:Z2) - version D; (f) reading rate (wpm) (E1:E2) - version D.

FIGURE 3: Inter-session agreement for the Z and E examiners using Bland and Altman method, showing mean differences between total errors, total time, and reading rate plotted against its means one week apart (Z1:Z2 and E1:E2) using the isiZulu paediatric rate of reading test on home language school children.

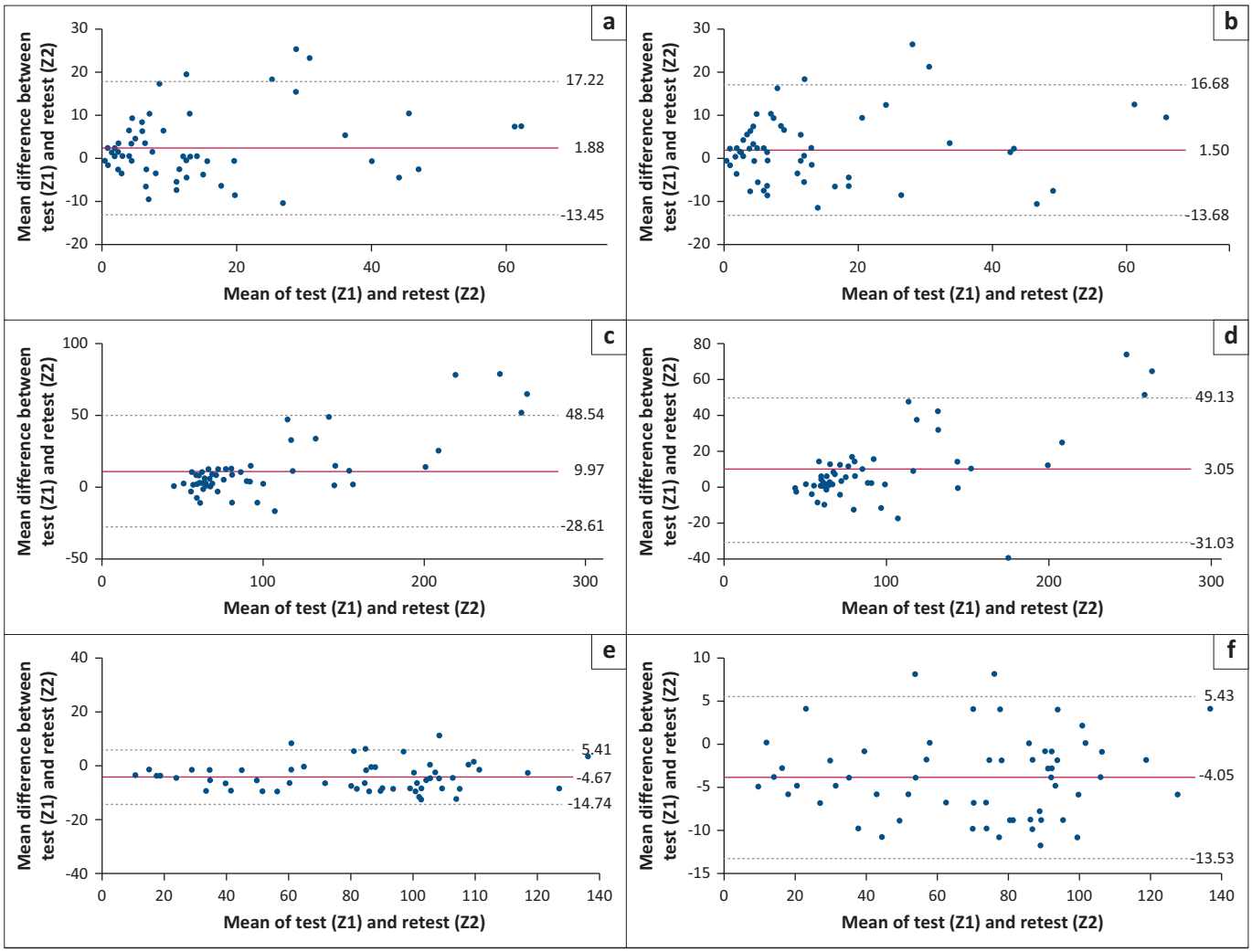
and E1:E2) at the HL and FAL schools are presented in Table 3, giving a measure of the inter-session reliability.

The ICC analysis across reading measures at isiZulu HL and FAL schools reveals consistently significant inter-session results between examiners with minimal variation as reflected by the high ICC values (most above 0.70) and narrow CIs (Table 3). Specifically, at the HL school, moderate associations were observed for total errors (≥ 0.664) and time (≥ 0.709) for both Z and E examiners, whereas at the FAL school, there was a good association for total errors (≥ 0.862) and an excellent association for time (≥ 0.922). Both Z and E examiners demonstrated excellent associations for reading rate (≥ 0.946) at both schools. Nevertheless, significant differences ($P < 0.05$) were detected in the inter-session means between examiners for all reading measures except for total errors at the FAL school ($P > 0.05$).

Furthermore, irrespective of the school context, the mean differences in test-re-test sessions for both examiners Z and E indicate higher values for total errors and time in session one compared to session two. Conversely, the reading rate showed higher values in session two.

Discussion

The isiZulu PRR Test was developed for primary school learners, grades one to seven, to be administered to isiZulu HL and FAL school children. The 2017 South African guidelines for compulsory education suggest that children should be between four and six years old for reception grade and seven years old for grade one.[28] Typically, grade seven is for students around 13 years old and marks the end of primary school. The age range of participants in this study however differs somewhat from these recommendations being from 6 years and 4 months to 14 years and 8 months, which includes Grades two to seven. This could suggest that younger participants (the 6-year-olds) possibly entered school earlier than recommended and older participants potentially experienced delays in school performance or enrolment. The study sample had a higher number of female participants compared to males,



Note: (a) Total errors (Z1:Z2) - version D; (b) total errors (E1:E2) - version D; (c) total time (s) (Z1:Z2) - version D; (d) votal time (s) (E1:E2) - version D; (e) reading rate in words per minute (wpm) (Z1:Z2) - version D; (f) reading rate (wpm) (E1:E2) - version D.

FIGURE 4: Inter-session agreement for the Z and E examiners using Bland and Altman method, showing mean differences between total errors, total time, and reading rate plotted against its means one week apart (Z1:Z2 and E1:E2) using the isiZulu paediatric rate of reading test on first additional language school children.

TABLE 3: The intraclass correlation coefficient and one-sample *t*-test for inter-session reliability of the Z and E examiners using version D of the isiZulu paediatric rate of reading test on home language and first additional language school children.

Schools	Examiners	Reading performance	Intraclass Correlation coefficient			One-sample *t*-test	
			ICC	Lower Bound	Upper Bound	Mean Difference	*P*-value
HL	Z1:Z2	Total errors	0.664	0.534	0.760	2.78	0.01
		Total time (s)	0.709	0.594	0.795	10.29	0.00
		Reading rate (wpm)	0.952	0.864	0.977	-5.75	0.00
	E1:E2	Total errors	0.671	0.539	0.769	3.06	0.040
		Total time (s)	0.712	0.599	0.797	11.67	0.025
		Reading rate (wpm)	0.949	0.855	0.976	-4.79	0.000
FAL	ZI:Z2	Total errors	0.862	0.778	0.916	1.88	0.670
		Total time (s)	0.923	0.837	0.960	9.97	0.000
		Reading rate (wpm)	0.975	0.852	0.991	-4.67	0.000
	E1:E2	Total errors	0.873	0.796	0.922	1.50	0.139
		Total time (s)	0.922	0.850	0.957	9.05	0.001
		Reading rate (wpm)	0.980	0.898	0.992	-4.05	0.000

ICC, intraclass correlation coefficient; FAL, first additional language; HL, home language; wpm, words-per-minute.

which could be attributed to greater interest among females in participating in the study as well as higher enrolment rates for females in schools.[29] In addition, South Africa has a larger female population relative to males,[30] which may contribute to the gender disparity observed in this study.

Inter-rater and inter-session reliability

Research defines reliability as the consistency and repeatability of test measurements. It reflects how reliable results of a test can be reproduced when the test is administered again under similar conditions but separated

by a time interval (inter-session reliability) or administered by different examiners (inter-rater reliability).[27] This concept ensures that a tool is standardised, providing accurate results free from measurement errors.[31]

The one-week interval between tests in this study aligns with practices from similar research conducted previously.[11] The ideal interval for reliability testing is between one day to one month.[32] This interval should be long enough to minimise memory effects, but short enough to avoid significant learning effects on test-re-test scores.[25] Given that children typically learn about seven new words daily[32] and that isiZulu PRR Test used random word placement per line; to counteract the impact of memorisation, a 1-week interval was deemed appropriate for assessing inter-session reliability in this study.

Inter-rater reliability of the Z and E examiners using the isiZulu paediatric rate of reading test

When multiple examiners are involved in administering and recording the results of a tool, the consistency between them is referred to as inter-rater reliability.[33] This study evaluated inter-rater reliability by comparing the results of total errors, time, and reading rate of the isiZulu (Z) and English-speaking (E) examiners on the same participants, aiming to determine if the isiZulu PRR Test could be used reliably by different examiners in the future who are not necessarily isiZulu-speaking. As the test was administered on two days (one week apart), inter-rater reliability was established on day one or session one (Z1:E1) and day seven or session two (Z2:E2) for both the isiZulu HL and FAL school children and will be discussed accordingly based on the Bland and Altman, ICC and one-sample *t*-test statistical analysis.

Some level of inconsistency was observed in the recording of total errors between the Z and E examiners reflected in the Bland-Altman plots (94%, 96/102 and 93%, 56/60) and one-sample *t*-test ($P = 0.027$) for session 1 at the FAL school. This variability could be related to the first-time exposure to the recording of errors and the level of experience between the examiners, with the E examiner having more exposure and experience than the Z examiner, hence contributing to the weaker inter-rater agreement for total errors. This inconsistency between examiners was similarly reported in a previous study using the MNREAD Chart, whereby more discrepancy was identified with the novice examiner.[34] This 'interpersonal dynamics' between examiners may negatively affect objectivity, which can be reduced by including examiners of equal expertise in reliability testing.[35] Furthermore, the considerable improvement in inter-rater reliability observed in session two at the HL school suggests that practice and experience make the results of the examiners more consistent and add value to the reliability of the test. A previous study reported that inter-rater reliability may be affected by the 'learning effect' or 'fatigue effect', with the former showing improvement in recordings between examiners in session two, while the latter reflects a decline.[35] It is therefore evident in the current study that the improvement can be attributed to the 'learning effect' between examiners. While this may reflect procedural familiarity between examiners, it is also possible that a learning effect occurred among the participants themselves, not only in terms of exposure to the reading text, but also with the testing format and expectations, potentially contributing to more consistent performances. However, as the design of the isiZulu PRR test consisted of random placement of words minimising memorisation of the test content, and that each participant was assessed by both examiners at the same time in the same session, keeping the testing procedure consistent, the impact of participant learning would least likely affect inter-rater reliability and have a greater influence on inter-session reliability or the overall validity of the test with the latter being included in a subsequent article. Following intense scrutiny of the Z and E examiners' scoresheets, variations in the recording of the type and number of errors between both examiners were observed, such as the recording of transposition as two substitutions, while a substitution may be recorded as one omission and one addition. In both instances, two errors may be recorded instead of one, which can also add to this discrepancy. Pre-test discussion and agreement on recording conventions such as the classification of errors between examiners may help eliminate such discrepancies. Finally, the subtle difference in the words 'yimi' for 'yini', and vice versa, may go unnoticed or be recorded as a substitution error, thereby affecting the number of errors recorded. A study performed on an isiZulu cohort revealed that subtle differences in spoken words are easily identifiable to a listener from the same linguistic background, in this case,[36] the Z examiner. A recommendation to future examiners is to ensure that the environment is free of distractions so that subtle differences are detected. Concepts of semantics, vocabulary, words, and language structure influence the reception of spoken words. While this may have influenced the recording of total errors in this study, the findings do not significantly affect the inter-rater reliability of the isiZulu PRR test. Further, as this study includes primary school children, a 95% CI may prove too stringent for this cohort hence requiring adjustments and clinical judgement.

The Bland-Altman plots, ICCs and one-sample *t*-test analysis for total time and reading rate display good inter-rater reliability, excellent linear relationship ($ICC \geq 0.986$) and no statistical differences between the Z and E examiners at the HL and FAL schools ($P > 0.05$), except for reading rate ($P = 0.007$) in session one at the FAL school. As the reading rate is inversely proportional to total error and total time,[11,24] the recording of the number of errors and time will affect the reading rate. With the recording of time being consistent between the examiners, the discrepancy then lies in the recording of errors. A careful inspection of the mean difference for total errors at the FAL school reflected that either the Z examiner recorded more errors, or the E examiner recorded fewer errors. This, in turn, may have contributed to a decrease in reading rate determined

by the Z examiner or an increase as determined by the E examiner. Together, these recordings may have had an additive effect on the reading rate results. As stated in a previous study, inter-rater reliability and agreement may be affected by the level of examiner training and hence should be factored into the process prior to data collection or usage of the test.[37]

Finally, the inter-rater results indicate that the reading measures are consistent, as the majority of them fall on and within the 95% LoA with *P*-values showing no statistically significant difference between the means of the examiners ($P > 0.05$), while the ICCs for all variables range from 0.848 to 1.000, suggesting good to excellent inter-rater reliability between the Z and E examiners when using the isiZulu PRR Test on isiZulu HL and FAL school participants.

Inter-session reliability of the optometric isiZulu paediatric rate of reading test for the Z and E examiners

Inter-session reliability involves evaluating the same participant multiple times over days, weeks, or months by the same examiner.[31,38,39] In this study, measurement was performed by the same examiner (Z or E) testing the same participant on day one (Z1 and E1) and one week later (Z2 and E2), using version D of the isiZulu PRR Test to assess total errors, time, and reading rate. Reliability is presented for each examiner as Z1:Z2 and E1:E2 and analysed using Bland-Altman plots, ICC, and one-sample *t*-test at isiZulu HL and FAL schools.

The inter-session averages for total errors and time obtained by both the examiners, at the HL and FAL schools, are higher in session one compared to session two (reflected by a positive value). This indicates that the participants made more errors, and took longer to read during session one, with performance improving in session two. This improvement is likely attributable to the participants' initial exposure to the test stimuli, alongside an adaptive learning effect characterised by increased familiarity with both the assessment items and the testing protocol. Such learning effects may have influenced the outcomes and should be acknowledged as a potential confounding factor when interpreting reading performance improvements across sessions. By incorporating a practice or familiarisation trial before formal testing, participants can become accustomed to the testing format, procedures and expectations, thereby reducing performance improvements attributable to procedural learning.

Furthermore, the Bland-Altman plots for the Z examiner only reflect that 94% and 93% of the inter-session results for total error lie within the LoA in the isiZulu HL and FAL schools, respectively, indicating inconsistencies with the recording done by the Z examiner as opposed to the E examiner. This too could be attributed to the first-time exposure of the Z examiner to the testing conditions, recording of results as well as clinical experience, whereby the E examiner had more years of experience in the administration of similar assessments and clinical practice. Such differences in expertise as well as communication style or procedural familiarity may therefore negatively affect objectivity.[35] As suggested for inter-rater reliability, stronger inter-session reliability can be expected with proper training of examiners before administration of the test.[35]

The inter-session Bland and Altman plots for time recorded by the Z and E examiners at the HL school, fell within the 95% LoA, but did not at the FAL school. The variation in total time could result from incorrect timing by the examiners either starting the timer too soon or stopping it too late, negatively impacting the inter-session recording of total time. Coincidentally with this trend observed for total time by both the Z and E examiners, this discrepancy in sessional results could be attributed to variation in participants' responses. Such variation influenced by inherent or environmental consequences, such as the development of growth during the period between measurements and an increase in vocabulary, has a considerable influence on the reliability of the results.[32,40] Furthermore, reading speed (time) has shown considerable variability over time impacting on reliability results.[41]

Finally, the mean difference for reading rate indicates a higher result for session two (negative value), supported by the lower number of errors and time in session two. Considering the inverse relationship between total error and time to reading rate,[11,24] with its compounding variation of both total errors and time in this session, this could have contributed to the increase in reading rate in session two. This slight inconsistency in the inter-session reading rate results is evident in the Bland-Altman plots at the HL school for both examiners (93% and 91% within the LoA), respectively. A possible explanation could be more exposure of the participants to the isiZulu language in the HL school than the FAL school during that week, which resulted in the improved reading rate observed in session two. In addition, participants' learning effect, following initial exposure, may have contributed to the improved performance and should be considered when interpreting future results. With the impact of developmental growth during the period between measurements, and the increase in vocabulary of a school child acquiring about seven new words a day,[32] hence impacting inter-session reliability, the application of the 95% LoA may not be ideal. Taffé et al. confirm that the width of the LoA for the Bland-Altman plots will apply only if the bias of the test and the measurement is constant with no variability during the entire testing procedure.[42] This therefore verifies that a 95% LoA may prove too stringent in the case of reading rate assessment in children hence the acceptability of the inter-session reliability of the isiZulu PRR test for reading rate measurement at the isiZulu HL school being interpreted as good.

While the *P*-values for all reading measures were statistically significant, reading rate values of less than 10 are considered to be clinically insignificant,[43,44] hence clinical relevance outweighs the statistical conclusion.[45] Finally, the ICCs for

total errors, time, and reading rate revealed moderate, moderate to excellent, and excellent inter-session reliability respectively at HL school, while good to excellent inter-session reliability for total errors, time, and reading rate for both examiners at the FAL school were observed. With the isiZulu PRR test showing inter-session reliability, it is suggested that before any conclusion could be drawn, repeated measurements are encouraged with a smaller interval rather than one week for true reading performance evaluation to reflect stronger reliability.

As noted, the potential for a learning effect among participants must be considered, particularly in repeated testing scenarios. While changing the interval between test and re-test is a common strategy, studies have shown that learning effects can still persist, especially in school-aged children who are rapidly developing cognitive and linguistic skills.[32] To mitigate this, standardised instructions and examiner behaviour, can help minimise procedural familiarity. Brief practice trials prior to the first test may also help stabilise early performance variability by allowing pupils to adjust to the testing process without contaminating the actual test data. Furthermore, limiting the visibility of scoring or feedback during testing may reduce behavioural adaptations. These strategies, combined with ongoing examiner training, can support more accurate and generalisable measures of reading performance in paediatric populations.

Conclusion

Clinical implications based on Version D of the isiZulu paediatric rate of reading test

The isiZulu PRR Test was designed with frequently used isiZulu words and optometric factors to assess reading rates in isiZulu-speaking primary school children with or without visual impairment. Analysis, including Bland-Altman plots, one-sample *t*-test, and ICCs, together with reference to clinical relevance, demonstrated good inter-rater and inter-session reliability for both isiZulu and English-speaking examiners, despite some variations, such as exposure and learning, which may have been encountered particularly in inter-session results. To mitigate such learning effects, common in repeated testing and especially in developing children, standardised instructions, examiner training, brief practice trials and limited feedback during testing can help ensure more consistent and valid results. Overall, the test is effective for evaluating and monitoring reading performance in children from isiZulu HL and FAL schools, supporting visual therapy, reading screening and managing visual impairment with reading devices. Its simplicity, quick administration, and low cost make it suitable for evaluating reading performance in South Africa, especially in rural areas with limited resources and overburdened district eye clinics and embrace the much-needed association between eyecare and education sectors, considering the impact of vision on learning.

Acknowledgements

This article includes content that overlaps with research originally conducted as part of Urvashni Nirghin's doctoral thesis titled, 'Development and validation of an isiZulu Paediatric Rate of Reading Test on South African isiZulu primary school children', submitted to the Department of Optometry, University of KwaZulu-Natal in 2019. The thesis was supervised by Rekha Hansraj and Lavanithum Joseph. Portions of the data, analysis, and/or discussion have been revised, updated and adapted for journal publication. The original thesis is publicly available at: https://researchspace.ukzn.ac.za/server/api/core/bitstreams/dff6f966-e45b-49f4-a679-fc4fcc3817f0/content. The author affirms that this submission complies with ethical standards for secondary publication and appropriate acknowledgement has been made of the original work.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

CRediT authorship contribution

Urvashni Nirghin: Data curation, Methodology, Analysis of results, Discussion, Writing – review & editing. Rekha Hansraj: Supervision. Lavanithum Joseph: Supervision. All authors reviewed the article, approved the final version for submission and publication, and take responsibility for the integrity of its findings.

Funding information

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability

The data that support the findings of this study are available from the corresponding author, Urvashni Nirghin upon reasonable request.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. The article does not necessarily reflect the official policy or position of any affiliated institution, funder, agency or that of the publisher. The authors are responsible for this article's results, findings and content.

References

1. Bailey I, Lovie J. The design and use of a new near-vision chart. Am J Optom Physiol Opt. 1980;57(6):378–387. https://doi.org/10.1097/00006324-198006000-00011
2. Legge G, Ross J, Luebker A, LaMays J. Psychophysics of reading. VII. The Minnesota low-vision reading test. Optom Vis Sci. 1989;66(12):843–853. https://doi.org/10.1097/00006324-198912000-00008

3. Ahn SJ, Legge GE, Luebker A. Printed cards for measuring low-vision reading Speed. Vis Res. 1995;35(13):1939–1944. https://doi.org/10.1016/0042-6989(94)00294-V
4. Wilkins A, Jeanes R, Pumfreyz P, 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(6):491–497. https://doi.org/10.1046/j.1475-1313.1996.96000282.x
5. Radner W, Willinger U, Obermayer W, Mudrich C, Velikay-Parel M, Eisenwort B. A new reading chart for simultaneous determination of reading vision and reading speed. Klin Monbl Augenheilkd. 1998;213(9):174–181. https://doi.org/10.1055/s-2008-1034969
6. Maaijwee K, Mulder P, Radner, W, Van Meurs JC. Reliability testing of the Dutch version of the Radner reading charts. Optom Vis Sci. 2008;85(5):353–358. https://doi.org/10.1097/OPX.0b013e31816bf58b
7. Alabdulkader B, Leat SJ. Toward developing a standardized Arabic continuous text reading chart. J Optom. 2017;10(2):84–94. https://doi.org/10.1016/j.optom.2016.03.003
8. Cheung JPY, Liu DSK, Lam CCC, Cheong AMY. Development and validation of a new Chinese reading chart for children. Ophthalmic Physiol Opt. 2015;35(5):514–521. https://doi.org/10.1111/opo.12228
9. Srinivasan K, Krishnan G, Wilkins A, Allen P. Reliability and validity of a Kannada rate of reading test. Indian J Ophthalmol. 2018;66(5):630–633. https://doi.org/10.4103/ijo.IJO_807_17
10. 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;43(4):629–639. https://doi.org/10.1111/opo.13134
11. Nirghin U. Design of a paediatric rate of reading test chart. Afr Vis Eye Heal. 2020;79(1):1–7. https://doi.org/10.4102/aveh.v79i1.536
12. Naidoo S, Gokool R, Ndebele H. The case of isiZulu for non-mother tongue speakers at the University of KwaZulu-Natal, South Africa – Is the compulsory language module promoting social cohesion? J Multiling Multicult Dev. 2018;39(4):356–368. https://doi.org/10.1080/01434632.2017.1393079
13. Howie S, Combrinck C, Tshele M, Roux K, Palane N, Mokoena G. PIRLS 2016: South African children's reading literacy achievement. Pretoria: Centre for Evaluation and Assessment, University of Pretoria; 2017.
14. Department of Basic Education. Annual report. Pretoria: Department of Basic Education; 2016.
15. Webb V. Language in South Africa: The role of language in national transformation, reconstruction and development. Illustrate. Amsterdam: John Benjamins; 2002, p. 356.
16. Song S, Levi D, Pelli D. A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. J Vis. 2014;14(5):1–37. https://doi.org/10.1167/14.5.3
17. Du Toit C. Towards a vocabulary for visual analysis: Using picture books to develop visual literacy with pre-service teachers. Mousaion. 2014;32(2):25–47. https://doi.org/10.25159/0027-2639/1688
18. Grosvenor T. Specification of visual acuity. In: Grosvenor T, editor. Primary care optometry. Boston, MA: Butterworth-Heinemann; 2002, p. 12–13.
19. Bennett A, Rabbetts R. Clinical visual optics. 4th ed. Rabbetts R, editor. Oxford: Butterworth-Heinemann; 2007, p. 1–451.
20. Koo TK, Li MY. Cracking the code: Providing insight into the fundamentals of research and evidence-based practice a guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15:155–163. https://doi.org/10.1016/j.jcm.2016.02.012
21. Kang H. Sample size determination and power analysis using the G*Power software. J Educ Eval Health Prof. 2021;18(17):1–12. https://doi.org/10.3352/jeehp.2021.18.17
22. Scheiman M, Wick B. Clinical management of binocular vision: Heterophoric, accommodative, and eye movement disorders. Philadelphia, PA: Lippincott Williams & Wilkins; 2014, p. 748.
23. Lailiyah N, Iswahyuni I, Lintangsari A. English spelling errors committed by the blind. Indones J Disabil Stud. 2020;7(1):66–71. https://doi.org/10.21776/ub.IJDS.2019.007.01.7
24. Deniz S, Yavuz M. Investigation of reading error types, reading levels and reading speeds of students with special learning difficulties. Cypriot J Educ Sci. 2020;15(4):686–698. https://doi.org/10.18844/cjes.v15i4.5051
25. Kimberlin CL, Winterstein G, Winterstein A. Validity and reliability of measurement instruments used in research. Am J Health Pharm. 2008;65:2276–2284. https://doi.org/10.2146/ajhp070364
26. Garcia-Martin E, Pinilla I, Idoipe M, Fuertes I, Pueyo V. Intra and interoperator reproducibility of retinal nerve fibre and macular thickness measurements using Cirrus Fourier-domain OCT. Acta Ophthalmol. 2011;89(1):23–30. https://doi.org/10.1111/j.1755-3768.2010.02045.x
27. Carey J. The critical reader-reliability. Am J Sports Med. 2017;45(14):3193–3195. https://doi.org/10.1177/0363546517743761
28. Lehohla P. Statistics South Africa. Education series volume III: Educational enrolment and achievement. Pretoria: Stats SA; 2016.
29. Martin V. Mid-year population estimates 2017. Pretoria: Stats SA; 2017.
30. Gadotti I, Vieira E, Magee D. Importance and clarification of measurement properties in rehabilitation. Rev Bras Fisioter. 2006;10(2):137–146. https://doi.org/10.1590/S1413-35552006000200002
31. Eduardo Paiva C, Marçon Barroso E, Cristina Carneseca E, et al. A critical analysis of test-retest reliability in instrument validation studies of cancer patients under palliative care: A systematic review. BMC Med Res Methodol. 2014;14(8):1–10. https://doi.org/10.1186/1471-2288-14-8
32. Snow C, Burns M, Griffin P. Preventing reading difficulties in young children. Washington, DC: National Academy Press; 1998, p. 416–465.
33. Hallgren K. Computing inter-rater reliability for observational data: An overview and tutorial. Tutor Quant Methods Psychol. 2012;8(1):23–34. https://doi.org/10.20982/tqmp.08.1.p023
34. Baskaran K, Macedo AF, He Y, Hernandez-Moreno L, Queirós T, Mansfield JS, Calabrèse A. Scoring reading parameters: An inter-rater reliability study using the MNREAD chart. PLoS One. 2019;14(6):e0216775. https://doi.org/10.1371/journal.pone.0216775
35. Belur J, Tompson L, Thornton A, Simon M. Interrater reliability in systematic review methodology: exploring variation in coder decision-making. Sociol Methods Res. 2021;50(2):837–865. https://doi.org/10.1177/0049124118799372
36. Panday P, Kathard H, Pillay M, Wilson W. The internal and external consistency of a speech reception threshold test for isiZulu speakers with normal hearing sensitivity. S Afr J Commun Disord. 2018;65(1):1–8. https://doi.org/10.4102/sajcd.v65i1.556
37. Shweta, Bajpai RC, Chaturvedi HK. Evaluation of inter-rater agreement and inter-rater reliability for observational data: An overview of concepts and methods. J Indian Acad Appl Psychol. 2015;41(3):20–27.
38. Onate J, Dewey T, Kollock R, et al. Real-time inter-session and interrater reliability of the functional movement screen. J Strength Cond Res. 2012;26(2):408–415. https://doi.org/10.1519/JSC.0b013e318220e6fa
39. Downs S, Marquez J, Chiarelli P. The Berg balance scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: A systematic review. J Physiother. 2013;59(2):93–99. https://doi.org/10.1016/S1836-9553(13)70161-9
40. Hu Y, Nesselroade J, Erbacher M, et al. Test reliability at the individual level. Struct Equ Model. 2016;23(4):532–543. https://doi.org/10.1080/10705511.2016.1148605
41. Brussee T, Van Nispen RMA, Van Rens GHMB. Measurement properties of continuous text reading performance tests. Ophthalmic Physiol Opt. 2014;34:636–657. https://doi.org/10.1111/opo.12158
42. Taffé P, Halfon P, Halfon M. A new statistical methodology overcame the defects of the Bland–Altman method. J Clin Epidemiol. 2020;124(2020):1–7. https://doi.org/10.1016/j.jclinepi.2020.03.018
43. Crawford L, Tindal G, Stieber S. Using oral reading rate to predict student performance on Statewide Achievement Tests. Educ Assess. 2001;7(4):303–323.
44. Kame'enui E, Fuchs L, Francis D, et al. The adequacy of tools for assessing reading competence: A framework and review. Educ Res. 2006;35(4):3–11. https://doi.org/10.3102/0013189X035004003
45. Heston TF. Statistical significance versus clinical relevance: A head-to-head comparison of the fragility index and relative risk index. Cureus. 2023;15(10):1–12. https://doi.org/10.7759/cureus.47741