Reading & Writing - Journal of the Literacy Association of South Africa
ISSN: (Online) 2308-1422, (Print) 2079-8245

The contributions of reading and phonological awareness for spelling in grade three isiXhosa learners

Authors:
Mikaela A. Daries[1,2]
Tracy N. Bowles[1,2]
Maxine N. Schaefer[3]

Affiliations:
[1]Department of Linguistics and Applied Language Studies, Faculty of Humanities, Rhodes University, Makhanda South Africa

[2]Centre for Social Development, Faculty of Education, Rhodes University, Makhanda, South Africa

[3]Department of Linguistics and Modern Languages, College of Human Sciences, University of South Africa, Pretoria, South Africa

Corresponding author:
Tracy Bowles,
t.bowles@ru.ac.za

Dates:
Received: 16 Feb. 2022
Accepted: 19 Aug. 2022
Published: 12 Dec. 2022

How to cite this article:
Daries, M.A., Bowles, T.N. & Schaefer, M.N., 2022, 'The contributions of reading and phonological awareness for spelling in grade three isiXhosa learners', *Reading & Writing* 13(1), a365. https://doi.org/10.4102/rw.v13i1.365

Background: One factor which is consistently highlighted in research on literacy is the lack of understanding of how literacy develops in the Southern-Bantu languages. In particular, little is known about spelling in the Southern-Bantu languages such as isiXhosa.

Objectives: Through the use of an initial exploratory study and a conceptual replication study, we examined the relationships between reading, phonological awareness, and spelling in isiXhosa grade 3 learners. The initial exploratory study sought to describe the relationships between reading and spelling, and phonological awareness and spelling in a sample of 49 grade 3 isiXhosa learners. We then conceptually replicated this study with a larger sample of 200 grade 3 isiXhosa learners. We expected that both reading and phonological awareness would be related to spelling and that the strength of the relationship between reading and spelling, and phonological awareness and spelling would vary with spelling ability, due to the changes that occur in the development of spelling.

Method: Cross-sectional, quantitative secondary data were used from two different projects to answer the research questions. Tasks of phonological awareness, oral reading fluency and spelling were developed and administered to the participants.

Results: We found that reading was a replicable predictor of spelling for grade 3 isiXhosa learners and that phonological awareness was influential only at the mid-range of spelling performance.

Conclusion: Our findings emphasise the importance of the reading – writing connection, and lend support for what has been found for other consistently written languages, adding to the growing body of knowledge of universal predictors of spelling development.

Keywords: spelling; phonological awareness; oral reading fluency; literacy; isiXhosa; conceptual replication.

Introduction

Reading and spelling are essential academic skills necessary for future educational success and successful participation in modern societies (Ehri 1987; Pretorius & Mokhwesana 2009; Zarić, Hasselhorn & Nagler 2020). It is, therefore, concerning that most South African learners cannot read, write, and comprehend at grade-appropriate levels, with many learners declared functionally illiterate (Spaull 2013). The 2016 Progress in International Literacy Study (PIRLS) re-emphasised this, revealing that 78% of South African grade 4 learners are not able to read for meaning (Howie et al. 2017). There are a number of contributory factors to South African learners' underachievement (De Vos, Van der Merwe & Van der Mescht 2015; Van der Berg et al. 2016). One factor, which is consistently highlighted, is the lack of understanding of how literacy develops in the alphabetically written Southern-Bantu languages (De Vos et al. 2015; Spaull, Pretorius & Mohohlwane 2020). In particular, little is known about spelling in the Southern-Bantu languages such as isiXhosa, with most studies focusing on decoding ability and the related role of phonological awareness (Makaure 2021; Schaefer, Probert & Rees 2020; Wilsenach 2013, 2019). Consequently, the predictors of spelling are not yet well understood in isiXhosa and further research is needed to enhance our understanding of spelling development in the Southern-Bantu languages.

In this article, the relationships between reading, phonological awareness, and spelling in isiXhosa were examined using secondary data from two studies. The first study was an exploratory cross-sectional study. The aim of the first study was to describe the relationships between reading and spelling, and phonological awareness and spelling in a small sample of grade 3 isiXhosa learners.

The first study was limited as it was cross-sectional and had a small sample size, giving us minimal evidence of the developmental relationship between variables, and leading to large confidence intervals of the estimates, respectively. We addressed these limitations by conducting a second, confirmatory study with four times the sample size.

The following two research questions were addressed in each study:

1. What are the unique contributions of reading and phonological awareness to spelling in grade 3 isiXhosa learners?
2. To what extent does the relationship between reading and spelling, and phonological awareness and spelling, differ on the level of spelling ability?

In this article we start by providing an overview of spelling development and its relation to reading and phonological awareness. We then address the methods, results and a brief discussion of each study respectively, highlighting the differences between study 1 and study 2 in the description of study 2. We end with a general discussion of our findings and provide a conclusion.

Literature review

Spelling development

Spelling is the ability to recognise and reproduce spoken language in a correct sequence of written symbols (Moats 2010; Mpiti 2012; Santoro, Coyne & Simmons 2006; Weiser & Mathes 2011). While there is not consensus on the development of spelling across languages, the main theories and perspectives are briefly reviewed below. These theories are, however, based mainly on English studies. Cross-linguistic studies on spelling development are relatively limited (Babayiğit 2009).

The earliest and most well-known model of spelling development was introduced by Gentry (1982), who suggested that learning to spell takes place in five stages, from reliance on phonological knowledge to a dependence on orthographic and morphological knowledge (Bourassa & Treiman 2010; Schlagel 2007). He posited that learners begin at the 'precommunicative stage' when the learner has partial knowledge of the alphabet but does not yet understand letter-sound relationships (Gentry 1982). At this stage, spellings are characterised by the strategy of randomly selecting letter strings to represent words (Ellis 1994). The second stage is the 'semi-phonetic stage', when learners have some grasp of letter– sound relationships, with spellings containing a partial mapping of phonetic content. It is at this stage that children often represent words, sounds or syllables with letters that match their letter names, for example, 'U' for 'you' (Ellis 1994). In the 'phonic' or third stage, they begin to use phonics to map letter-sound correspondences, for example, 'haws' for 'house' (Ellis 1994). In the later stages (transitional and correct stages), learners move from phonological spelling and start to incorporate morphological and orthographic knowledge when spelling (Gentry 1982).

Critiques of the stage theory have stated that it provides a linear account of spelling and oversimplifies the picture with little consideration of the linguistic characteristics of the language being learnt (Bourassa & Treiman 2010). For example, different languages have specific morphological and orthographic knowledge needed to be acquired by learners and therefore some skills are grasped more easily than others and at different periods in spelling development (Bourassa & Treiman 2010). Nevertheless, the theory acknowledges the developmental shifts when learning to spell, especially from a reliance on phonological knowledge to a dependence on orthographic and morphological knowledge (Bourassa & Treiman 2010; Ravid 2012; Schlagal 1989). However, Babayiğit (2009) argued that children may utilise both phonological and orthographic spelling processes in the early stages of development depending on the demand of the task, their level of alphabetic knowledge, phonological awareness and print experience.

Cross-linguistic studies which have examined spelling development and the specific skills most relevant for spelling development across languages are growing in number but most studies have focused on English (Caravolas 2004; Georgiou et al. 2020). Nevertheless, there is a general acknowledgment that spelling relies on letter-sound knowledge (e.g. Finnish, Greek and English: Georgiou et al. 2012), phonological awareness (e.g. English, Spanish, Slovak, and Czech: Caravolas et al. 2012), morphological awareness (e.g. English: Apel et al. 2012), orthographic knowledge (e.g. English: Apel 2009) and reading experience (e.g. Finnish: Leppänen et al. 2006; Lerkkanen et al. 2004; French, Dutch, German and Greek; Georgiou et al. 2020). As noted above, phonological awareness and reading are related to spelling ability. The interrelation between these concepts are addressed in more detail in the next sections.

The relationship between phonological awareness and spelling

Phonological awareness is defined as the awareness of the sound structure of a language and the ability to perceive or manipulate sounds in a linguistic sequence (Moll et al. 2014). Phonological awareness is typically divided into the following linguistic components: onset-rime, syllables and phonemes (Anthony & Francis 2005; Chard & Dickson 1999; Stahl & Murray 1994; Treiman 1991).

There is an extensive body of research which confirms that phonological awareness is a significant predictor of both reading and spelling across different languages (Adams 1990; Babayiğit & Stainthorp 2007; Bryant et al. 1990; Caravolas, Volín & Hulme 2005; Hulme & Snowling 2015; Landerl & Wimmer 2008; Zarić et al. 2020). Notably it has been revealed that phonemic awareness is the strongest predictor of early word reading in alphabetic languages (Melby-Lervåg, Lyster & Hulme 2012). Similarly, research on the Southern-Bantu languages converge on the importance of phonological awareness for reading (Lekgoko & Winskel 2008; Malda, Nel & Van De Vijver 2014; Wilsenach 2013,

2015, 2019). However, the relationship between spelling and phonological awareness in the Southern-Bantu languages is less clear.

International research reports that phonological awareness plays a concurrent and longitudinal role in spelling performance (Caravolas et al. 2012; Harrison et al. 2016; Landerl & Wimmer 2008; Lervåg & Hulme 2010; Veber Nielsen & Juul 2016). In particular, research has suggested that phonological awareness and spelling should have a stronger relationship in languages with transparent orthographies due to the consistent nature of the phoneme to grapheme and grapheme to phoneme mappings (e.g. Leppänen et al. 2006; Öney & Durgunoglu 1997). Phonological awareness is a necessary prerequisite for the conversion of phonemes to graphemes, the primary requirement of spelling across all alphabetically written languages. However, as noted in Georgiou et al. (2012), these assertions are not made in consideration of the individual characteristics of languages with transparent orthographies.

De Bree and Van den Boer's (2019) study of Dutch spelling in grades 1 and 2 found that phonological awareness significantly predicted spelling in regression analyses. However, when reading was added into their regression model, phonological awareness along with other contributors (visual attention span and rapid automatised naming) were no longer significant, confirming the importance of reading for early spelling (De Bree & Van den Boer 2019; Desimoni, Scalisi & Orsolini 2012; Keuning & Verhoeven 2008; Leppänen et al. 2006) and reaffirming the link between phonological awareness and reading (De Bree & Van den Boer 2019; Frith 1985). De Bree and van den Boer (2019) suggest that phonological awareness constitutes mainly as an underlying skill of reading and that phonological awareness relates to spelling but that it is not necessarily sufficient in explaining how learners convert phonemes to graphemes. Thus, they suggest that further research is needed to investigate specific correlates and underlying skills related to spelling, for example, vocabulary knowledge, morphological awareness, and orthographic learning (De Bree & Van den Boer 2019; Kim, Apel & Al Otaiba 2013). These results are of particular significance as Dutch, like isiXhosa, is considered to have a transparent orthography in which the regularity of the orthographic system is higher in grapheme–phoneme relations (forward regularity, e.g. reading,) than it is in phoneme–grapheme relations (backward regularity, e.g. encoding) (Bosman, Vonk & Van Zwam 2006; Daries & Probert 2020). For example, in isiXhosa, there are a number of complex graphemes for which multiple letters represent single sounds e.g. <dl>, <ph>, <tsh> (Spaull et al. 2020).

Southern African research on spelling in languages other than English confirms the relationship between phonological awareness and spelling. Phonological awareness was associated with spelling ability in Herero for a sample of grade 2 to grade 5 learners (Veii 2003), in Oshikwanyama for a sample of grade 1 learners, assessed again in grade 2 (Nghikembua 2020), and in isiXhosa for a small sample of grade 3 isiXhosa learners (Diemer 2015). Makaure (2021) found similar results with 134 Northern Sotho-English bilingual learners. The study's findings showed that phonological skill was one of the best predictors of Northern Sotho spelling. Common to all these studies was the need to develop tests of phonological awareness suitable to each language, and sample. These were real word phonological awareness tests of phoneme isolation using sound matching (e.g. Which one is the odd one out: *sava*, *sina*, *zuva*? (Veii 2003)), real-word phoneme isolation tasks (e.g. What is the first sound of *pedu*? [Nghikembua 2020]), and pseudoword tasks of segmenting (e.g. Break up this word into its sounds: *gofotsa*) and elision (Say *hlenama* without */hl/* [Diemer 2015]). Although various tasks have been used, an effect of phonological awareness on spelling was found across different South African languages and grade levels.

The relationship between reading and spelling

Reading can be understood as the starting point of spelling, as what learners are able to read can provide the basis for what they are able to spell (Bear et al. 2004). When learners read or practise reading, they are recognising common letter patterns in words, mapping sounds onto letters and letter patterns which develops their orthographic representations of words necessary for spelling (Conrad 2008). This suggests that spelling can provide an indication of learners' reading ability. Furthermore, Gentry (2004:11) states that 'learners learn to read by spelling' and knowledge of the alphabetic principle necessary for spelling supports reading. This is because spelling and reading rely on the same lexical representations (Ehri 1997, 2005; Perfetti 1997; Templeton 2004).

While the exact nature of the relationship between reading and spelling is not clearly defined and understood, there are a number of studies which indicate strong and significant correlations between spelling and reading ability (Abbott et al. 2010; Caravolas, Hulme & Snowling 2001; De Bree & Van den Boer 2019; Ehri 1997; Foorman et al. 1991; Georgiou et al. 2020; Ritchey 2008). Whilst there are few studies which examine the effect of reading on spelling in the African languages found in Southern Africa, existing studies confirm that reading ability is associated with spelling. For example, Nghikembua (2020) reported a significant association between grade 1 reading and grade 2 spelling using a path analysis model in Oshikwanyama. Given cross-linguistic evidence on the role of reading for spelling in orthographies which vary in transparency (e.g. Georgiou et al. 2020), one might expect to find a positive association between reading ability and spelling in other African languages.

Study 1

Method

Study 1 was exploratory. We used cross-sectional data first reported by Daries and Probert (2020).

Participants

The participants were 49 (20 female, 29 male; M_{age} = 9.1 years, SD_{age} = 0.7 years) grade 3 children attending one Quintile 3 isiXhosa- medium of instruction school in a small low-income township of the Eastern Cape of South Africa. Participants were assessed in the third term of 2019.

Measures

Tasks of phonological awareness, oral reading fluency and spelling were developed and administered to the participants. The phonological awareness and spelling instruments used in both studies can be found on the Open Science Framework (OSF) (https://osf.io/cesgf/).

Spelling: The spelling task consisted of 12 real words which ranged from two to four syllables in length (Daries & Probert 2020). As described in Daries and Probert (2020), a number of linguistic decisions relating to word frequency, word length in syllables and grapheme complexity, were made in the design of the spelling task. The spelling task was scored using a binary set of codes, that is, correct and incorrect per item. The item scores were summed (max = 12) for each learner. The reliability of the task was 0.95 as measured by Cronbach's alpha (Daries & Probert 2020).

Reading: Reading ability was measured using a one minute timed oral reading fluency task. Oral reading fluency is defined as the ability to read a text quickly, accurately, and with meaningful expression (Fuchs et al. 2001; Spaull et al. 2020). Oral reading fluency is often used as a measure of learners' reading ability, and is traditionally calculated using a words-correct-per-minute (WCPM) score which reflects both accuracy and speed (Spaull et al. 2020).

The oral reading fluency task required the learners to read aloud an isiXhosa short story from a grade 3 book for one minute. The story *Iyho! Mkhulu lo mnqathe! (Wow! A gigantic carrot)* was used from the Vula Bula books and was 102 words long with a mean of 6.8 words per sentence, and 7.4 letters per orthographic word. The errors made during the tasks were recorded and subtracted from the total words the learner attempted to read to provide a WCPM score. The WCPM score was used in the analysis.

Phonological awareness: Phonological awareness was measured using 48 pseudoword items (12 items per condition) at the phoneme and syllable levels using elision and isolation tasks (Diemer 2015; Probert 2016). Pseudowords, ranging from two to four syllables long, rather than real words were used to reduce the influence of semantic information on responses. These tasks were developed in consideration of the linguistic and cognitive complexity appropriate for isiXhosa and for grade 3 learners (Diemer, Van der Merwe & De Vos 2015). The item scores per task were summed per learner and these total scores were used in the principal components analysis to derive a final score. The reliability of the overall task was 0.92 as measured by Cronbach's alpha.

Procedures

The order of task administration was fixed. The phonological awareness and oral reading fluency tasks were administered individually before the spelling task, which was administered in a group setting. Each word of the spelling task was read aloud twice for dictation. Task presentation for phonological awareness was fixed so that phoneme subtasks (isolation then elision) were presented first, followed by the syllable subtasks (isolation then elision).

Data analysis procedures: We used R version 4.04 (R Core Team 2021) for the analysis. The code is available on OSF (https://osf.io/cesgf/). A principal components analysis, using the psych package (Revelle 2021), was run on the four phonological awareness indicators to reduce the number of variables in the later analysis. We first specified four factors (one for each measure). There was only one eigenvalue above one, indicating that the data fit a one-principal component model. We extracted the scores of the first principal component for later analysis and labelled this phonological awareness. We used the raw total score for spelling to make it more comparable across studies.

To estimate the predictors of spelling at the mean of spelling (research question 1), we used hierarchical linear regression analyses. We ran a series of models first with only the intercept, then added the covariate age, followed by reading, and phonological awareness in the final step so that we could estimate the unique effects of phonological awareness and reading ability, then calculated the change in R^2 at each step. The residuals were not normally distributed so we used 2000 bootstrap resamples to estimate the slopes and their confidence intervals.

To determine if the slope of phonological awareness for spelling, and the slope of reading for spelling differed by spelling ability level, we fitted quantile regression models using the quantreg package (Koenker 2021) in R. Quantile regression allows researchers to determine whether the relationship between predictors and an outcome variable depend on the value of the outcome variable, such as whether the outcome variable is low or high (Petscher & Logan 2014). For example, quantile regression was used by James et al. (2021) to determine whether morphological awareness had a similar contribution to reading comprehension for low-, middle- and high-ability students on reading comprehension.

The data for study 1 come from a cross-sectional design. To interpret the results of the quantile regression, we argue that students with low scores on the spelling task are at a different developmental level than students who score higher on the task. We interpret the various total scores on the spelling task as reflecting performance that corresponds to different spelling ability levels. Thus, we fit the quantile regression to the data at the 10th, 50th and 90th quantiles (as done by James et al. 2021; McIlraith & Reading Research Consortium 2018) corresponding to earlier spelling

development, typical spelling development for the sample, and advanced spelling ability, respectively. We determined that quantile regression would best allow us to address the second research question.

The interpretation of a quantile model slopes differs slightly from that of linear regression (Petscher & Logan 2014). It is recommended to use centred variables for the predictors such that zero represents mean ability. At each quantile, then, the intercept corresponds to the value of the outcome at that quantile. The slope of the predictor (let us call it x) should be interpreted as the difference in the outcome between a participant with mean of x and a participant at 1 SD on x (Petscher & Logan 2014). The confidence interval indicates whether this difference (i.e. slope term) is different from zero. In this study, we performed 1000 bootstrap resamples to estimate the confidence intervals. Quantile comparisons are used to statistically test whether the slopes are different at each quantile (Petscher & Logan 2021). A goodness of fit statistic is also calculated for each quantile (Koenker & Machado 1999).

Results

The descriptive statistics and correlations among the variables are presented in Table 1. Visualisations of the spread of the raw data, and assumptions of the linear regression, are presented in the OSF project https://osf.io/cesgf/. Phonological awareness and reading had moderate to strong correlations with spelling (*r* above 0.65). Age was weakly correlated with spelling and reading, and had a weak negative correlation with phonological awareness (r = -0.35, p = 0.04). We categorised the sample into reading fluency groups using the suggested isiXhosa fluency benchmarks for children in grade 3 (Ardington et al. 2021, see Table A1 in the appendix). In our sample, 31 children (61%) were not reading at the lower fluency threshold of 20 WCPM. 17 children (33%) read within the fluency zone of 20–34 WCPM which was a similar proportion found for isiXhosa learners in term 1 of grade 3 (Ardington et al. 2021). Only one child read above the upper threshold of 35 or more WCPM. The sample includes slower readers than those sampled by Ardington et al. (2021).

The results of the multiple linear regression and quantile regression are presented in Table 2, following, and in Figure 1 in the next section. The results were different for each estimation method. At the mean, both reading and phonological awareness were significant predictors accounting for 39.4% and 14.7% of the variance, respectively (see Linear Regression in Table 2). The confidence intervals for reading and phonological awareness overlapped one another, indicating that the slopes may not be different from one another. The trend in the quantile regression results was that the difference in ability on spelling was larger for a participant 1 SD higher on reading (slope = 2.36), than phonological awareness (slope = 1.33) at the 10th quantile, and the trend was reversed at the 50th quantile (reading slope = 0.74, phonological awareness slope = 2.07). However, there was no difference in performance for participants at the mean or 1 SD above the mean on the predictors at the 90th quantile, possibly because the participants at the 90th quantile were perfoming at ceiling level on the spelling task. A quantile comparison test indicated that the slopes did not differ across quantiles, which is confirmed also by the overlap in confidence intervals at each quantile.

Discussion

Our first study was exploratory, and examined to what extent phonological awareness and reading predicted spelling at the mean, and at different quantiles of spelling performance, corresponding to low, medium and high performance on spelling. With reference to the first research question, we found that at the mean (linear regression), both reading and phonological awareness were significant (unique) predictors of spelling. With reference to the second research question, the quantile regression presented evidence that the strength of the predictors may differ at different levels of spelling ability, with reading ability leading to larger differences in spelling at the 10th quantile, and phonological awareness

TABLE 1: Study 1: Descriptive statistics and Pearson correlations for phonological awareness, reading, spelling and age.

Variable (unit/maximum score)	1. Age (years)	2. Phon. eli. (/12)	3. Phon. iso. (/12)	4. Syl. eli. (/12)	5. Syl. iso. (/12)	6. PA: PCA	7. Reading (wcpm)	8. Spelling (/12)
Descriptive statistics (N = 49)								
Mean	9.1	4.0	5.1	4.6	5.6	0.0	15.8	9.3
Median	9.0	2.0	5.0	5.0	6.0	0.0	16.0	11.0
SD	0.7	4.3	3.5	3.2	4.1	1.0	10.5	4.0
Min	8.0	0.0	0.0	0.0	0.0	-1.7	0.0	0.0
Max	10.0	12.0	12.0	11.0	12.0	2.2	41.0	12.0
Pearson correlations								
2. Phon. elision	-0.35*	-	-	-	-	-	-	-
3. Phon. isolation	-0.25	0.50***	-	-	-	-	-	-
4. Syl. elision	-0.22	0.53***	0.41**	-	-	-	-	-
5. Syl. isolation	-0.03	0.19	0.35*	0.47**	-	-	-	-
6. PA PCA	-0.29*	0.76***	0.76***	0.82***	0.64***	-	-	-
7. Reading	-0.19	0.48**	0.52***	0.58***	0.36*	0.66***	-	-
8. Spelling	-0.21	0.54***	0.53***	0.62***	0.49***	0.73***	0.66***	-

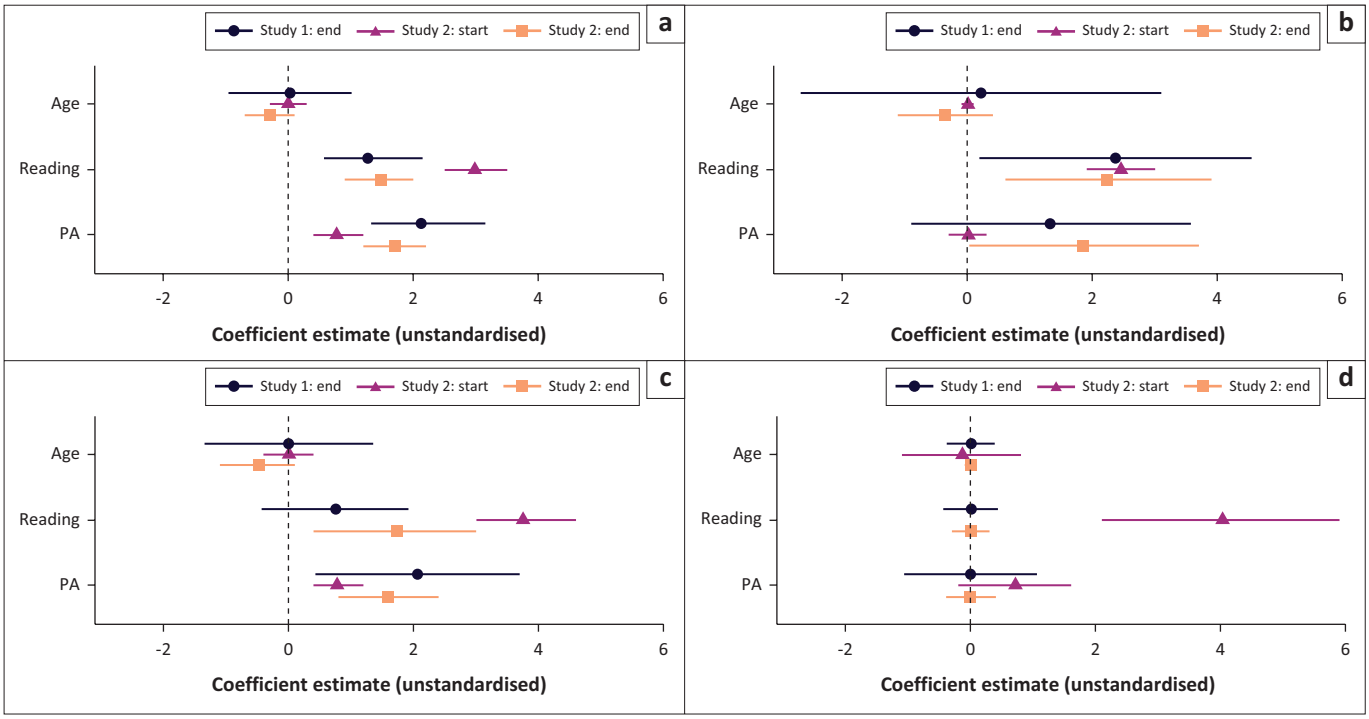
Note: *p < 0.05, **p < 0.01, ***p < 0.001.

Phon., Phoneme; Eli., Elision; Iso., Isolation; Syl., Syllable; PA, Phonological Awareness; PCA, Principal Components Analysis; wcpm, words-correct-per-minute.

TABLE 2: Study 1: Results of linear and quantile regressions of spelling predicted by age, reading and phonological awareness.

Model	Parameter	Estimate	SE	95% CI	*t*	*p*	ΔR^2	Goodness of fit
Linear regression								
Final model	(Intercept)	8.99	4.56	-0.10, 17.91	1.78	0.080	-	-
	Age	0.03	0.50	-0.96, 1.01	0.06	0.950	0.02	-
	Reading	1.27	0.40	0.57, 2.14	2.50	**0.017**	0.39	-
	PA	2.11	0.47	1.32, 3.14	4.10	**< 0.001**	0.15	-
Quantile regression								
QR-10	(Intercept)	3.75	13.02	-22.48, 29.97	0.29	0.775	-	0.54
	Age	0.21	1.43	-2.68, 3.10	0.15	0.884	-	-
	Reading	2.36	1.08	0.19, 4.54	2.20	**0.034**	-	-
	PA	1.33	1.11	-0.91, 3.57	1.20	0.237	-	-
QR-50	(Intercept)	9.80	6.03	-2.35, 21.95	1.62	0.111	-	0.25
	Age	0.00	0.67	-1.35, 1.35	0.00	0.999	-	-
	Reading	0.74	0.58	-0.44, 1.91	1.26	0.213	-	-
	PA	2.07	0.81	0.43, 3.70	2.55	**0.014**	-	-
QR-90	(Intercept)	12.00	1.75	8.48, 15.52	6.90	**< 0.001**	-	0.00
	Age	0.00	0.19	-0.38, 0.38	0.00	1.000	-	-
	Reading	0.00	0.22	-0.44, 0.44	0.00	1.000	-	-
	PA	0.00	0.53	-1.06, 1.06	0.00	1.000	-	-

Note: Linear regression: F (3, 45) = 21.62, p < 0.001; N = 49; nested models compared to calculate ΔR^2. Quantile regression: N = 49; Goodness of fit provided for the model at each quantile and not per predictor. QR-10, quantile regression at 10th quantile; QR-50, quantile regression at 50th (median) quantile; QR-90, quantile regression at 90th quantile.



Note: PA, phonological awareness.

FIGURE 1: Model coefficients with 95% confidence intervals for linear and quantile regressions for study 1 (end of Grade 3) and study 2 (start and end of Grade 3). (a) Linear Regression; (b) Quantile Regression: tau = 0.10; (c) Quantile Regression: tau = 0.50; (d) Quantile Regression: tau = 0.90.

ability leading to larger differences in spelling ability at the 50th quantile. At the 90th quantile, the participants were at ceiling levels of performance in spelling. However, given the small sample size, all confidence intervals were wide, leading to the slope differences not reaching significance.

Study 1 provided evidence that both phonological awareness and reading are significant predictors of real word spelling in this sample, but there was insufficient precision to address whether these relationships do differ by levels of spelling ability.

Study 2

Study 2 was conducted to determine whether the results of study 1 were replicable, and is, therefore, a confirmatory study. Based on the results of study 1, and a review of the literature, we hypothesised that we would find similar trends to study 1 in study 2.[1] In study 2 we expected that both reading and phonological awareness would be positively associated with

1. These hypotheses and how they would be tested were pre-registered before the research commenced. The pre-registration is accessible on OSF (https://osf.io/gbqx6) and also provides the results of the power analysis undertaken for the linear regression analysis.

spelling for grade 3 isiXhosa learners as phonological awareness is a necessary prerequisite for the conversion of phonemes to graphemes, the primary requirement of spelling, across all alphabetically written languages (Georgiou et al. 2012). Additionally, research has attested to a relationship between spelling and reading (Abbott et al. 2010; Caravolas et al. 2001; De Bree & Van den Boer 2019; Ehri 1997; Foorman et al. 1991; Ritchey 2008). In addressing the first research question, our null hypothesis (H[0]1) was that there would be no relationship between reading, phonological awareness and spelling, and the alternate hypothesis (H[1]1) was that reading and phonological awareness would be positively associated with spelling.

Furthermore, we expected that the strength of the relationship between phonological awareness and spelling, and reading and spelling would vary with spelling ability, due to the changes that occur in the development of spelling. At early stages of spelling, reading should play a larger role, with a shift to phonemic decomposition as children become better spellers and then a shift to orthographic and morphemic spelling as children become more accurate spellers (Ellis 1994; Frith 1985; Gentry 1982).

The second null hypothesis (H(0)2) was that the strength of the relationship between reading and spelling, and phonological awareness and spelling, would not differ by spelling ability. We had specific alternate hypotheses for each level of spelling ability tested in our models, based on the results of study 1:

> H(1)2a: Reading will have a stronger positive association with spelling, compared to phonological awareness, at the lower end of spelling ability.
>
> H(1)2b: Phonological awareness will have a stronger positive association with spelling, compared to reading, at the middle of the range of spelling ability.
>
> H(1)2c: Reading and phonological awareness will not be associated with spelling at the higher end of spelling ability.

Method

We used secondary data from a collaborative literacy intervention project to perform a conceptual replication of study 1. The intervention project focuses on strengthening the isiXhosa literacy skills of children in grades 1 to 3 in five schools in low socioeconomic areas in the Eastern Cape of South Africa, as well as the upskilling and capacity building of teachers and literacy ambassadors (classroom assistants). Study 2 draws participants from the same population of grade 3 learners in low-income Quintile 3 township schools in the Eastern Cape. All participants attended schools in which the intervention took place i.e. there was no comparison/control group. The intervention results are not of interest to this article and will not be considered further.

Study 2 differs notably from study 1 in the following ways: phonological awareness was measured using a shorter task, the data was from an intervention project, participants were assessed at the start and end of grade 3, and the participants were sampled in 2021 which is the second year that COVID-19 impacted the quantity and quality of education in South Africa.

Participants

Two hundred participants (104 female, 96 male) were sampled from 12 classrooms (1–28 participants per classroom) and five schools (15–75 learners per school). The participants' self-reported age at the start of grade 3 was 8.5 years (SD = 1 year) on average. The age difference between the study 1 and study 2 participants represented a large effect, $t(247) = 3.97$, $p < 0.001$, $d = 0.63$, 95% CI [0.31, 0.95]. The study 1 sample reported their age later in the school year (term 3) than the study 2 sample (term 1), so this difference somewhat corresponds to the 7.5 month expected difference in average ages. Thus, in terms of school socioeconomic status and self-reported age, the participants were similar between studies.

Measures

Study 2 also included tasks of phonological awareness, oral reading fluency and spelling, with some differences addressed here.

Spelling: The spelling task administered in study 2 used eight of the 12 real words used in study 1. Four words were replaced to enhance the complexity of the task as some ceiling effects were identified with the grade 3 sample used in Daries and Probert (2020). The Cronbach's alpha internal consistency for this task was 0.86.

Reading: Reading was measured in the same way as study 1, with a new text. The story titled *Uhobe noMbovane* [Ant and Dove] was used. This reading passage was 72 words long with a mean of 5.1 words per sentence, and 7.8 letters per orthographic word.

Phonological awareness: Phonological awareness was measured slightly differently in study 2 with regards to the stimuli, number of items and the way the final score was calculated. The stimuli were 12 real words on which participants were asked to perform the same operations as in study 1, i.e. phoneme and syllable elision and isolation. Because of the reduced number of items, Cronbach's alpha was lower ($\alpha = 0.70$). Only a total sum score was available in the dataset. The phonological awareness tasks in both studies measured the same construct so their effect on the outcome variable (spelling) can be compared with the following caveats: raw or z scores cannot be directly compared as the tasks are not identical, and the lower reliability of the task in study 2 and the way the final score was calculated means that the final score includes more error compared to true score than in study 1. The effect of the latter is that the relationship between phonological awareness and the outcome will be attenuated (Plonsky & Derrick 2016). We return to the effect of this difference when interpreting the results of this conceptual replication.

Procedures

The data collection procedures were the same as in study 1, except that data were collected at the start and end of grade 3, whereas for study 1, data were only collected towards the end of grade 3. The dataset was missing 3.1% of data points but under listwise deletion we would lose 69 (34.5%) participants. Age and spelling at the start of grade 3 and spelling at the end of grade 3 were the variables with the most missing data. Participants who had at least one missing data point had lower reading and phonological awareness scores (small effect size) and lower spelling scores (medium effect size) at the start of grade 3 than participants with complete data, i.e., the data were Missing at Random. Since listwise deletion would drastically reduce the sample size and bias the estimates (Woods et al. 2021), we imputed the missing data using predictive mean matching in mice (multiple imputation by chained equations) (Van Buuren & Groothuis-Oudshoorn 2011). We imputed 10 datasets with 10 iterations each and saved the final model's final iteration as the dataset we worked with.

The data analysis was similar to study 1, except that we used mixed models to account for the clustering of the data at school or teacher level. We used lme4 (Bates et al. 2015), lmerTest (Kuznetsova, Brockhoff & Christensen 2017), and lqmm (Geraci 2014; Geraci & Bottai 2014) to estimate these models. Using mixed models, we are able to calculate the intra-class correlation (ICC) to determine the proportion of variance in the outcome explained by the clustering variable (classroom or school). Because we can account for clustering, the standard errors associated with the fixed effects (i.e. variables of interest: age, reading and phonological awareness) are less biased than using analytic methods which do not account for the clustering (Theobald 2018).

Ethical considerations

Ethical clearance was obtained from the Rhodes University Ethical Standards Committee for both study 1 (RU-HSD--2019-0461-452) and study 2 (RU-HSD- 2020-1195-3307). Ethical clearance was obtained from the Eastern Cape Department of Basic Education for each study.

Results

The descriptive statistics and correlations among the variables are presented in Table 3. Visualisations of the spread of the raw data, and assumptions of the linear regression, are presented in the OSF project https:// osf.io/cesgf/. On average, the oral reading fluency of the group was low at the start of grade 3 (M = 12.7, SD = 13.6), and had improved by the end of grade 3 (M = 30.7, SD = 19). There was a large-effect size difference in the reading abilities of study 2 participants at the end of grade 3 compared to study 1, $t(247) = -5.29$, $p < 0.001$, $d = -0.84$, 95% CI [-1.16, -0.52]. There were more learners at or above the threshold of 20 WCPM in study 2 (70%) than in study 1 (37%) at the same point in the academic year. There was no difference for spelling between the studies at the same point in the academic year, $t(247) = 0.94$, $p = 0.348$, $d = 0.15$, 95% CI [-0.16, 0.46]. The phonological awareness scores could not be compared directly because of the different way the construct was measured in each study. The correlations among variables was moderate to strong, which was similar to study 1. The correlation between phonological awareness, and reading and spelling was slightly lower than in study 1, probably due to the lower reliability of the phonological awareness task in study 2. As in study 1, age was negatively and weakly correlated with phonological awareness at the end of grade 3. Additionally, the same negative correlation was found for spelling at the end of grade 3 in study 2.

TABLE 3: Study 2: Descriptive statistics, effect-size difference (Cohen's *d*) to study 1 and Pearson correlations for phonological awareness, reading, spelling and age.

Variable	T1: Start of Grade 3				T2: End of Grade 3		
(unit/maximum score)	1. Age (years)	2. Reading (wcpm)	3. PA (/12)	4. Spelling (/12)	5. Reading (wcpm)	6. PA (/12)	7. Spelling (/12)
Descriptive statistics (N = 200)							
Mean	8.5	12.7	5.5	5.3	30.7	8.7	8.7
Median	8.0	9.0	5.0	5.0	31.5	10.0	11.0
SD	1.0	13.6	3.0	4.3	19.0	3.0	4.0
Min	6.0	0.0	0.0	0.0	0.0	0.0	0.0
Max	12.0	70.0	12.0	12.0	72.0	12.0	12.0
Difference to study 1							
t	4.0	1.5	-	5.6	-5.3	-	0.9
p	**< 0.001**	0.138	-	**< 0.001**	**< 0.001**	-	0.348
Cohen's *d* [95% CI]	0.63 [0.31, 0.95]	0.2 [-0.1, 0.6]	-	0.9 [0.6, 1.2]	-0.8 [-1.2, -0.5]	-	0.2 [-0.2, 0.5]
Pearson correlations							
2. T1 Reading	-0.18*	-	-	-	-	-	-
3. T1 PA	-0.08	0.52***	-	-	-	-	-
4. T1 Spelling	-0.12	0.78***	0.53***	-	-	-	-
5. T2 Reading	-0.13	0.69***	0.46***	0.66***	-	-	-
6. T2 PA	-0.21**	0.40***	0.43***	0.46***	0.54***	-	-
7. T2 Spelling	-0.21**	0.55***	0.54***	0.67***	0.60***	0.61***	-

Note: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

PA, phonological awareness; PA tasks not directly comparable between study 1 and study 2.

The results of the linear mixed model and linear quantile models are presented in Table 4 and Figure 1. First, the intra class correlation (ICC) was examined to determine whether school or teacher random effects should be included in the model. The ICC for school level was 0.161 at the start of grade 3 and 0.025 at the end of grade 3, and at teacher level the ICC was 0.084 at the start of grade 3 and 0.032 at the end of grade 3. We decided to use school random effects at the start of grade 3 and teacher random effects at the end of grade 3 because these levels affected the spelling outcome the most. We then fitted maximal models with varying intercepts, varying intercepts and slopes (uncorrelated) and varying intercepts and slopes (correlated). Only the varying intercepts models converged and were used for further analysis. In the final step, we determined the fixed effects by comparing a null model (only the cluster predictor) to successive models adding age, then reading, then phonological awareness. For both timepoints, the model fit improved with the addition of each fixed effect. Thus, we report the models using varying intercepts for school (start of grade 3) and teacher (end of grade 3) with all three predictors, in Table 4.

The linear mixed model, which estimated the mean, found that at both the start and end of grade 3, reading and phonological awareness were significant predictors of spelling. The 95% CI for reading at the end of grade 3 (beta = 1.47, [0.9, 2.0]) overlapped with the estimate in study 1 (beta = 1.27, [0.57, 2.14]. The 95% CI for phonological awareness for study 2 (beta = 1.71, [1.2, 2.2]) and study 1 (beta = 2.11, [1.32, 3.14] also overlapped, despite the change in measurement. As expected, the width of the 95% CI was smaller in study 2 because of the larger sample. The 95% CIs of the start of grade 3 estimates (reading beta = 2.97, [2.5, 3.5], phonological awareness beta = 0.76, [0.4, 1.2]) did not overlap with those at the end of grade 3 for either study 1 or 2; the reading estimate was much larger, and the phonological awareness estimate much smaller.

The linear mixed quantile regression results indicated that participants who had 1 SD higher ability in reading and phonological awareness had a similar increment in spelling score compared to participants at mean of these variables at the 10th and 50th quantiles at the end of grade 3 (Table 4). At the 90th quantile, neither were significant, replicating what was found in study 1. The clustering at teacher level accounted for a large proportion of the variance at the 90th quantile (ICC = 0.294) compared to the other quantiles (ICC = 0.016_{10th}, ICC_{50th} = 0.023). The start of grade 3 results established reading as a significant predictor at all three

TABLE 4: Study 2: Results from mixed-effects models at the mean, and 0.1, 0.5 and 0.9 quantiles at the start of grade 3 (school random effects) and end of grade 3 (teacher random effects).

Model	Parameter	T1: Start of Grade 3					T2: End of Grade 3				
		Est.	SE	95% CI	*t*	*p*	Est.	SE	95% CI	*t*	*p*
Linear mixed model											
	ICC	0.161	-	-	-	-	0.032	-	-	-	-
	(Intercept)	5.28	0.26	4.8, 5.9	20.3	0.000	8.74	0.33	8.1, 9.4	26.8	0.000
	Age	0.01	0.20	-0.4, 0.4	0.0	0.970	-0.31	0.21	-0.7, 0.1	-1.5	0.146
	Reading	2.97	0.23	2.5, 3.5	13.0	**< 0.001**	1.47	0.25	0.9, 2.0	5.8	**0.000**
	PA	0.76	0.23	0.4, 1.2	3.3	**0.001**	1.71	0.25	1.2, 2.2	6.7	**0.000**
	RE (Intercept) SD	0.37	-	0, 0.9	-	-	0.84	-	0, 1.4	-	-
	RE Residual SD	2.67	-	2.4, 3.0	-	-	2.83	-	2.5, 3.1	-	-
Linear mixed quantile model											
QR-10	ICC	0.093	-	-	-	-	0.016	-	-	-	-
	(Intercept)	5.17	0.35	4.5, 5.9	-	< 0.001	5.44	0.84	3.8, 7.1	-	< 0.001
	Age	0.00	0.07	-0.1, 0.1	-	0.999	-0.36	0.37	-1.1, 0.4	-	0.335
	Reading	2.46	0.27	1.9, 3.0	-	**< 0.001**	2.23	0.77	0.6, 3.9	-	0.006
	PA	0.00	0.13	-0.3, 0.3	-	0.999	1.85	0.89	0.04, 3.7	-	0.043
	RE (Intercept) SD	1.22	-	-	-	-	0.64	-	-	-	-
	RE Residual SD	3.82	-	-	-	-	5.04	-	-	-	-
QR-50	ICC	0.008	-	-	-	-	0.023	-	-	-	-
	(Intercept)	4.96	0.27	4.4, 5.5	-	< 0.001	8.79	0.42	7.8, 9.8		< 0.001
	Age	0.00	0.21	-0.4, 0.4	-	0.999	-0.49	0.31	-1.1, 0.1		0.121
	Reading	3.75	0.39	3.0, 4.6	-	**< 0.001**	1.73	0.65	0.4, 3.0		**0.010**
	PA	0.78	0.21	0.4, 1.2	-	**< 0.001**	1.58	0.43	0.8, 2.4		**< 0.001**
	RE (Intercept) SD	0.26	-	-	-	-	0.48	-	-	-	-
	RE Residual SD	2.80	-	-	-	-	3.07	-	-	-	-
QR-90	ICC	0.014	-	-	-	-	0.294	-	-	-	-
	(Intercept)	8.52	1.18	6.2, 10.9	-	< 0.001	12.00	0.63	11.0, 13.3	-	< 0.001
	Age	-0.14	0.49	-1.1, 0.8	-	0.768	0.00	0.07	-0.1, 0.1	-	0.999
	Reading	4.02	0.94	2.1, 5.9	-	**< 0.001**	0.00	0.18	-0.3, 0.3	-	0.999
	PA	0.71	0.43	-0.2, 1.6	-	0.103	0.00	0.15	-0.4, 0.4	-	0.999
	RE (Intercept) SD	0.56	-	-	-	-	2.13	-	-	-	-
	RE Residual SD	4.80	-	-	-	-	3.30	-	-	-	-

Note: ICC, intra-class correlation; PA, phonological awareness; RE, random effects.

tested quantiles. Participants with 1 SD higher on reading had higher scores on spelling than those at the mean at each quantile, with a slightly smaller difference at the 10th quantile. Participants who differed by 1 SD on phonological awareness only had higher scores on spelling at the 50th quantile. The school that a participant was in had the largest effect at the 10th quantile (ICC = 0.093), than the other two quantiles (ICC_{50th} = 0.008, ICC_{90th} = 0.014).

Discussion

In the second study, we aimed to replicate the findings of study 1. We found evidence in support of our first hypothesis that reading and phonological awareness are related to spelling ability. These findings were replicated at the mean of spelling where the 95% CIs of the estimates overlapped for study 1 and study 2. We also find evidence in support of our second hypothesis. Comparing the start and end of grade 3 linear mixed-model results in study 2, we found that reading was a stronger predictor of spelling at the lower ability level (i.e. start of grade 3) than at the higher spelling ability level (i.e. end of grade 3). The mixed-quantile regression results partially supported H(1)2a and H(1)2c. At the lower end of spelling ability, 1 SD change in reading lead to larger gains in spelling compared to 1 SD change in phonological awareness. When spelling ability was at ceiling level, neither phonological awareness nor reading were significant predictors. We had hypothesised that phonological awareness would become a stronger predictor of spelling at the mid-range of spelling ability. Although it did reach significance at the 50th quantile at the start of grade 3, and the 10th and 50th quantiles for the end of grade 3, a 1 SD change in phonological awareness compared to someone at the mean lead to an increase in spelling scores that was lower than those of similar magnitude to a 1 SD difference on reading.

General discussion

In the two studies, we examined the relationships between reading, phonological awareness and spelling for isiXhosa grade 3 learners. Research has confirmed the importance of both reading and phonological awareness for spelling, and in this study, we contribute to the growing body of evidence from languages other than English by using data from isiXhosa. Additionally, we included a conceptual replication of our own work to determine the replicability of our findings. After running our analyses on a small sample of participants at the end of grade 3, we sought to replicate the analyses using a sample four times larger, and assessed at both the start and end of grade 3. The participants were similar in terms of the schools they attended and their mean age extrapolated to the end of grade 3. The participants' end of grade 3 spelling scores did not differ between studies, but the study 2 participants had higher reading abilities at the end of grade 3. Thus, in most respects, the participants were similar.

These two studies indicated that in isiXhosa, reading and phonological awareness make significant unique predictions to spelling (research question 1). Reading was a consistent positive concurrent predictor of spelling in grade 3. This supports the findings of Nghikembua (2020), who reported a significant association between grade 1 reading and grade 2 spelling in Oshikwanyama. Further this result seems to be in line with the perspective of Bear et al. (2004) who explains that reading is the starting point of spelling in that what learners are able to read provides the basis for what they are capable of spelling. When learners read or practise reading they are recognising common letter patterns in words, mapping sounds onto letters and letter patterns which develops their orthographic representations of words necessary for spelling (Conrad 2008).

Phonological awareness was a consistent positive predictor of spelling at the mean of spelling across studies, which is consistent with the findings for Herero grade 2 to grade 5 learners (Veii 2003), Oshikwanyama grade 1 and grade 2 learners (Nghikembua 2020) and isiXhosa grade 3 learners (Diemer 2015). However, its relationship to spelling differed across the spelling ability distribution (research question 2). In the quantile models, phonological awareness was significant at the mid-range of spelling ability (i.e. when the intercept corresponded to raw scores of 5 to 10). In the replication (study 2), phonological awareness's effect was weaker than reading. This result is similar to what de Bree and van den Boer (2019) found for transparently written Dutch. The difference in the strength of the relationship between phonological awareness and spelling at the 50th quantile in study 1 and study 2 may be due to how phonological awareness was measured in each study, and also possibly due to the study 2 participants having higher overall reading skills at the end of grade 3. Overall, these two studies provide evidence for Ellis' (1994) observation that young spellers begin with early attempts at copying whole words (which relies on reading experience) to developing degrees of mastery of the alphabetic principle by using phonetic decomposition (which relies on phonological awareness) to finally develop orthographic and morphemic spelling (which relies on orthographic and morphological awareness) (Frith 1985; Gentry 1982). The influence of phonological awareness was restricted to the middle of the range of performance. At ceiling levels of spelling performance, neither phonological awareness nor reading were significant predictors, highlighting that other skills, such as orthographic or morphological awareness, may play a role.

At the method level, we undertook a conceptual replication of study 1. We did this to address the imprecision in our estimates due to the small sample size, a limitation of many applied linguistics research articles. Notably the measurement of phonological awareness in each study differed. The measurement of phonological awareness in study 2 more closely represented how phonological awareness has been measured in the past: a sum score of fewer than 15 items. When summing all items, these items are given equal weight in the phonological awareness construct. Reliability is lower in a shorter than a longer test. However, in study 1, a principal components analysis was used on almost 50 items. In principal components analysis, items are weighted based on

their contribution to the overall phonological awareness construct, thus providing a 'truer' representation of the phonological awareness construct. Additionally, longer tests are more reliable. The briefer nature of the phonological awareness task included as a sum score in study 2, compared to a more detailed task converted to a principal components score in study 1 may have led to an underestimation of the role of phonological awareness in study 2. This is a limitation applicable across studies which use sum scores, and tasks with reliability on the lower end.

Overall, the findings of our study confirm the association found between reading and spelling ability, and phonological awareness and spelling ability in transparent orthographies using cross-sectional data. The strengths of our study included a replication aspect in which similar participants as those in study 1 were sampled at the start and end of grade 3. Future studies could sample participants before they begin formal instruction and longitudinally follow their spelling development until grade 3. Longitudinal studies can best address how various predictors relate to spelling over the course of development. We did not include additional variables known to affect spelling such as orthographic or morphological awareness (Apel 2009; Apel et al. 2012), and we included only one indicator variable per construct. Future research could include multiple measures of the same construct and use a latent variable framework to better account for measurement error. Sum scores are easy to use in practice and to interpret but the sole use of sum scores could lead to an underestimation of effects. As far as possible, researchers are encouraged to share item-level data, not only total summed scores, in their open access datasets which would allow these kinds of latent variable analyses.

Our findings emphasise the importance of the reading–writing connection. Spelling was most related to reading ability, and phonological awareness was associated with spelling at the mid-range of performance. These results suggest that teachers should make sure children know their letter–sound correspondences, can blend these to form words, and can use these skills automatically for fluent reading which may in turn support spelling development.

In conclusion, through the use of two studies, we found that reading was a replicable predictor of spelling for grade 3 isiXhosa learners and that phonological awareness was influential only at the mid-range of spelling performance. These findings support what has been found for other consistently written languages, adding to the growing body of knowledge of universal predictors of reading and spelling development.

Acknowledgements

The authors wish to extend their sincere thanks to the schools and learners for their generous participation in study 1 and study 2. Further, a note of thanks must be given to the Rhodes University Centre for Social Development as the implementing partner of the literacy intervention, funded by the Volkswagen Community Trust, for allowing us access to the secondary data used in study 2. We also acknowledge James et al. (2021) for making their code available (https://osf.io/96542/), which we were able to use for the current analyses.

Competing interests

The first and second authors are part of the research team on the literacy intervention from which the secondary data were accessed for study 2. However, the authors have declared that no competing interests exist.

Authors' contributions

M.D. contributed to the conceptualisation, methodology, writing – original draft, writing, review and editing, visualisation. T.B. contributed to the conceptualisation, writing – original draft, writing-review and editing, supervision. M.S. contributed to the conceptualisation, analysis, writing review and editing, visualisation.

Funding information

Study 1 received funding from the Rhodes University Research Community Grant Fund. The Volkswagen Community Trust funded the project from which the secondary data was used for study 2.

Data availability

The data are available via request from the corresponding author. The code for study 1 and study 2, and (links to) research instruments are included on the Open Science Framework (https://osf.io/cesgf/). An earlier draft of this article with associated preregistration are available at https://edarxiv.org/4pkbe/ and https://osf.io/gbqx6, respectively.

Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any affiliated agency of the authors.

References

Abbott, R.D., Berninger, V.W. & Fayol, M., 2010, 'Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7', *Journal of Educational Psychology* 102(2), 281–298. https://doi.org/10.1037/a0019318

Adams, M.J., 1990, *Beginning to read: Thinking and learning about print*, MIT Press, Cambridge, MA.

Anthony, J.L. & Francis, D.J., 2005, 'Development of phonological awareness', *Current Directions in Psychological Science* 14(5), 255–259. https://doi.org/10.1111/j.0963-7214.2005.00376.x

Apel, K., 2009, 'The acquisition of mental orthographic representations for reading and spelling development', *Communication Disorders Quarterly* 31(1), 42–52. https://doi.org/10.1177/1525740108325553

Apel, K., Wilson-Fowler, E.B., Brimo, D. & Perrin, N.A. 2012, 'Metalinguistic contributions to reading and spelling in second and third grade students', *Reading and Writing* 25(6), 1283–1305. https://doi.org/10.1007/s11145-011-9317-8

Ardington, C., Wills, G., Pretorius, E., Mohohlwane, N. & Menendez, A., 2021, 'Benchmarking oral reading fluency in the early grades in Nguni languages', *International Journal of Educational Development* 84, 102433. https://doi.org/10.1016/j.ijedudev.2021.102433

Babayiğit, S., 2009, 'Reading and spelling in transparent alphabetic orthographies: Points of convergence and divergence', in C. Wood & V. Connelly (eds.), *Contemporary perspectives on reading and spelling*, pp. 133–148, Routledge, London.

Babayiğit, S. & Stainthorp, R., 2007, 'Preliterate phonological awareness and early literacy skills in Turkish', *Journal of Research in Reading* 30(4), 394–413. https://doi.org/10.1111/j.1467-9817.2007.00350.x

Bates, D., Mächler, M., Bolker, B. & Walker, S., 2015, 'Fitting linear mixed-effects models using {lme4}', *Journal of Statistical Software* 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bear, D.R., Invernizzi, M., Templeton, S. & Johnston, F., 2004, *Words their way*, Merrill Prentice Hall, Upper Saddle River, NJ.

Bosman, A., Vonk, W. & Van Zwam, M., 2006, 'Spelling consistency affects reading in young Dutch readers with and without dyslexia', *Annals of Dyslexia* 56(2), 271–300. https://doi.org/10.1007/s11881-006-0012-4

Bourassa, D.C. & Treiman, R., 2010, 'Linguistic foundations of spelling development', in Wyse D, Andrews R, Hoffman J, editors. *The Routledge international handbook of English, language and literacy teaching*, pp. 206–216, Routledge, London

Bryant, P.E., MacLean, M., Bradley, L.L. & Crossland, J., 1990, 'Rhyme and alliteration, phoneme detection, and learning to read', *Developmental Psychology* 26(3), 429–438.

Caravolas, M., 2004, 'Spelling development in alphabetic writing systems', *European Psychologist* 9(1), 3–14. https://doi.org/10.1027/1016-9040.9.1.3

Caravolas, M., Hulme, C. & Snowling, M.J., 2001, 'The foundations of spelling ability: Evidence from a 3-year longitudinal study', *Journal of Memory and Language* 45(4), 751–774. https://doi.org/10.1006/jmla.2000.2785

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E. et al., 2012, 'Common patterns of prediction of literacy development in different alphabetic orthographies', *Psychological Science* 23(6), 678–686. https://doi.org/10.1177/0956797611434536

Caravolas, M., Volín, J. & Hulme, C., 2005, 'Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children', *Journal of Experimental Child Psychology* 92(2), 107–139. https://doi.org/10.1016/j.jecp.2005.04.003

Chard, D.J. & Dickson, S.V., 1999, 'Phonological awareness: Instructional and assessment guidelines', *Intervention in School and Clinic* 34(5), 261–270. https://doi.org/10.1177/105345129903400502

Conrad, N.J., 2008, 'From reading to spelling and spelling to reading', *Journal of Educational Psychology* 100(4), 869–878. https://doi.org/10.1037/a0012544

Daries, M.A. & Probert, T.N., 2020, 'A linguistic analysis of spelling errors in grade 3 isiXhosa home-language learners', *Reading & Writing* 11(1). e1–e10. https://doi.org/10.4102/rw.v11i1.265

De Bree, E. & Van den Boer, M., 2019, 'Knowing what we don't know: Cognitive correlates of early spelling of different target types', *Reading & Writing* 32(8), 2125–2148. https://doi.org/10.1007/s11145-019-09936-9

De Vos, M., Van der Merwe, K. & Van der Mescht, C., 2015, 'A linguistic research programme for reading in African languages to underpin CAPS', *Journal for Language Teaching* 48(2), 148–177. https://doi.org/10.4314/jlt.v48i2.7

Desimoni, M., Scalisi, T.G. & Orsolini, M., 2012, 'Predictive and concurrent relations between literacy skills in grades 1 and 3: A longitudinal study of Italian children', *Learning and Instruction* 22(5), 340–353. https://doi.org/10.1016/j.learninstruc.2012.02.002

Diemer, M.N., 2015, 'The contributions of phonological awareness and naming speed to the reading fluency, accuracy, comprehension and spelling of grade 3 isiXhosa readers', unpublished MA dissertation, Rhodes University, Grahamstown.

Diemer, M.N., Van der Merwe, K. & De Vos, M., 2015, 'The development of phonological awareness literacy measures for isiXhosa', *Southern African Linguistics and Applied Language Studies* 33(3), 325–334. https://doi.org/10.2989/16073614.2015.1108769

Ehri, L.C., 1987, 'Learning to spell and read words', *Journal of Reading Behaviour* 19(1), 5–31.

Ehri, L.C., 1997, 'Learning to read and learning to spell are one and the same, almost', in C. Perfetti, L. Rieben & M. Fayol (eds.), *Learning to spell: Research, theory and practice*, pp. 237–269, Erlbaum, Mahwah, NJ.

Ehri, L.C., 2005, 'Learning to read words: Theory, findings, and issues', *Scientific Studies of Reading* 9(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4

Ellis, N.C., 1994, 'Longitudinal studies of spelling development', in G.D.A. Brown & N.C. Ellis (eds.), *Handbook of spelling: Theory, process and intervention*, pp. 155–190, Wiley, Chichester.

Foorman, B.R., Francis, D.J., Novy, D.M. & Liberman, D., 1991, 'How letter-sound instruction mediates progress in first-grade reading and spelling', *Journal of Educational Psychology* 83(4), 456–469. https://doi.org/10.1037/0022-0663.83.4.456

Frith, U., 1985, 'Beneath the surface of developmental dyslexia', in K. Patterson, J. Marshall & M. Coltheart (eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading*, pp. 301–330, Erlbaum, London.

Fuchs, L.S., Fuchs, D., Hosp, M.K. & Jenkins, J.R., 2001, 'Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis', *Scientific Studies of Reading* 5(3), 239–256. https://doi.org/10.1207/S1532799XSSR0503_3

Gentry, J.R., 1982, 'Analysis of developmental spelling skills in GNYS AT WRK', *The Reading Teacher* 36(2), 192–200.

Gentry, J.R., 2004, *The science of spelling: The explicit specifics that make great readers and writers (and spellers!)*, Heinemann, Portsmouth.

Georgiou, G.K., Torppa, M., Landerl, K., Desrochers, A., Manolitsis, G., De Jong, P.F. & Parrila, R., 2020, 'Reading and spelling development across languages varying in orthographic consistency: Do their paths cross?', *Child Development* 91(2), e266–e279. https://doi.org/10.1111/cdev.13218

Georgiou, G.K., Torppa, M., Manolitsis, G., Lyytinen, H. & Parrila, R., 2012, 'Longitudinal predictors of reading and spelling across languages varying in orthographic consistency', *Reading and Writing: An Interdisciplinary Journal* 25(2), 321–346. https://doi.org/10.1007/s11145-010-9271-x

Geraci, M., 2014, 'Linear quantile mixed models: The lqmm package for Laplace quantile regression', *Journal of Statistical Software* 57(13), 1–29. https://doi.org/10.18637/jss.v057.i13

Geraci, M. & Bottai, M., 2014, 'Linear quantile mixed models', *Statistics and Computing* 24(3), 461–479. https://doi.org/10.1007/s11222-013-9381-9

Harrison, G., Goegan, L.D., Jalbert, R., McManus, K., Sinclair, K. & Spurling, J., 2016, 'Predictors of spelling and writing skills in first- and second-language learners', *Reading and Writing: An Interdisciplinary Journal* 29(1), 69–89. https://doi.org/10.1007/s11145-015-9580-1

Howie, S., Combrinck, C., Roux, K., Tshele, M., Mokoena, G. & McLeod Palane, N., 2017, *PIRLS 2016 progress in international reading literacy study 2016: South African children's reading literacy achievement*, Centre for Evaluation and Assessment, Pretoria, viewed 01 March 2021, from http://www.sajce.co.za.

Hulme, C. & Snowling, M.J., 2015, 'Learning to read: What we know and what we need to understand better', *Child Development Perspectives* 1(7), 1–5. https://doi.org/10.1111/cdep.12005

James, E., Currie, N.K., Tong, S.X. & Cain, K., 2021, 'The relations between morphological awareness and reading comprehension in beginner readers to young adolescents', *Journal of Research in Reading* 44(1), 110–130. https://doi.org/10.1111/1467-9817.12316

Keuning, J. & Verhoeven, L., 2008, 'Spelling development throughout the elementary grades: The Dutch case', *Learning and Individual Differences* 18(4), 459–470. https://doi.org/10.1016/j.lindif.2007.12.001

Kim, Y.S., Apel, K. & Al Otaiba, S., 2013, 'The relation of linguistic awareness and vocabulary to word reading and spelling for first-grade students participating in response to intervention', *Language, Speech, and Hearing Services in Schools* 44(4), 337–347. https://doi. org/10.1044/0161-1461(2013/12-0013)

Koenker, R., 2021, *quantreg: Quantile Regression. R package version 5.86*, viewed 10 July 2021, from https://CRAN.R-project.org/package=quantreg.

Koenker, R. & Machado, J.A.F., 1999, 'Quantile regression, inference processes and goodness of fit', *Journal of the American Statistical Association* 94(448), 1296–1310.

Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B., 2017, 'lmerTest package: Tests in linear mixed effects models', *Journal of Statistical Software* 82(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Landerl, K. & Wimmer, H., 2008, 'Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up', *Journal of Educational Psychology* 100(1), 150–161. https://doi.org/10.1037/0022-0663.100.1.150

Lekgoko, O. & Winskel, H., 2008, 'Learning to read Setswana and English: Cross-language transference of letter knowledge, phonological awareness and word reading skills', *Perspectives in Education* 26(4), 57–73, viewed 1 June 2021, from https://hdl.handle.net/10520/EJC87497.

Leppänen, U., Aunola, K., Niemi, P. & Nurmi, J.E., 2008, 'Letter knowledge predicts grade 4 reading fluency and reading comprehension', *Learning and Instruction* 18(6), 548–564. https://doi.org/10.1016/j.learninstruc.2007.11.004

Lerkkanen, M., Rasku-Puttonen, H., Aunola, K. & Nurmi, J., 2004, 'The developmental dynamics of literacy skills during the first grade', *Educational Psychology* 24(6), 793–810. https://doi.org/10.1080/0144341042000271782

Lervåg, A. & Hulme, C., 2010, 'Predicting the growth of early spelling skills: Are there heterogeneous developmental trajectories?', *Scientific Studies of Reading* 14(6), 485–513. https://doi.org/10.1080/10888431003623488

Makaure, Z.P., 2021, 'The contribution of phonological processing skills to early literacy development in northern Sotho-English bilingual children: A longitudinal investigation', doctoral dissertation.University of South Africa

Malda, M., Nel, C. & Van De Vijver, F.J., 2014, 'The road to reading for South African learners: The role of orthographic depth', *Learning and Individual Differences* 30, 34–45. https://doi.org/10.1016/j.lindif.2013.11.008

McIlraith, A.L. & Language and Reading Research Consortium, 2018, 'Predicting word reading ability: A quantile regression study', *Journal of Research in Reading* 41(1), 79–96. https://doi.org/10.1111/1467-9817.12089

Melby-Lervåg, M., Lyster, S.-A.H. & Hulme, C., 2012, 'Phonological skills and their role in learning to read: A meta-analytic review', *Psychological Bulletin* 138(2), 322–352. https://doi.org/10.1037/a0026744

Moats, L.C., 2010, *Speech to print: Language essentials for teachers*, Brookes, Baltimore, MD.

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N. et al., 2014, 'Cognitive mechanisms underlying reading and spelling development in five European orthographies', *Learning and Instruction* 29, 65–77. https://doi.org/10.1016/j.learninstruc.2013.09.003

Mpiti, T., 2012, 'The nature of spelling errors of grade three isiXhosa background learners in English first additional language', Doctoral dissertation, University of Fort Hare.

Nghikembua, T.N., 2020, *Learning to read and spell in Oshikwanyama language precursors, dynamics and teacher knowledge of early literacy instruction*, University of Jyväskylä, viewed 22 May 2021, from http://urn.fi/URN:ISBN:978-951-39-8385-7.

Öney, B. & Durgunoglu, A.Y., 1997, 'Beginning to read in Turkish: A phonologically transparent orthography', *Applied Psycholinguistics* 18(1), 1–15. https://doi.org/10.1017/S014271640000984X

Perfetti, C.A., 1997, 'The psycholinguistics of spelling and reading', in C.A. Perfetti, L. Rieben & M. Fayol (eds.), *Learning to spell: Research, theory, and practice across languages*, pp. 21–38, Lawrence Erlbaum Associates Publishers, Mahwah, NJ.

Petscher, Y. & Logan, J.A.R., 2014, 'Quantile regression in the study of developmental sciences', *Child Development* 85(3), 861–881. https://doi.org/10.1111/cdev.12190

Plonsky, L., & Derrick, D. J., 2016, 'A Meta-Analysis of reliability coefficients in second language research', *The Modern Language Journal*, 100 (2), 539-553. https://doi.org/10.1111/modl.12335

Pretorius, E., & Mokhwesana, M., 2009, 'Putting reading in Northern Sotho on track in the early years: Changing resources, expectations and practices in a High Poverty School', *South African Journal of African Languages* 29(1), 54–73. https://doi.org/10.1080/02572117.2009.10587317

Probert, T., 2016, 'A comparative study of syllables and morphemes as literacy processing units in word recognition: IsiXhosa and Setswana', unpublished MA dissertation, Rhodes University, Grahamstown.

R Core Team, 2021, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, viewed 1 July 2021, from https://www.R-project.org/.

Ravid, D.D., 2012, 'The psycholinguistics of spelling: Phonology and beyond', in *Spelling morphology*, pp. 21–39, Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-0588-8_2

Revelle, W., 2021, *psych: Procedures for personality and psychological research*, North-Western University, Evanston, IL, viewed 10 July 2021, from https://CRAN.R-project.org/package=psychVersion=2.1.6.

Ritchey, K., 2008, 'The building blocks of writing: Learning to write letters and spell words', *Reading and Writing: An Interdisciplinary Journal* 21(1–2), 27–47. https://doi.org/10.1007/s11145-007-9063-0

Santoro, L.E., Coyne, M.D. & Simmons, D.C., 2006, 'The reading–spelling connection: Developing and evaluating a beginning spelling intervention for children at risk of reading disability', *Learning Disabilities Research and Practice* 21(2), 122–133. https://doi.org/10.1111/j.1540-5826.2006.00212.x

Schaefer, M., Probert, T. & Rees, S., 2020, 'The roles of phonological awareness, rapid automatised naming and morphological awareness in isiXhosa', *Per Linguam* 36(1), 90–111. https://doi.org/10.5785/36-1-878

Schlagal, R.C., 1989, 'Constancy and change in spelling development', *Reading Psychology: An International Quarterly* 10(3), 207–232. https://doi.org/10.1080/0270271890100302

Schlagel, B., 2007, *Best practices in spelling and handwriting*, in S Graham, C. A. MacArthur, & J. Fitzgerald (eds.), 'Best practices in writing instruction (pp. 179–201), The Guilford Press, New York.

Spaull, N., 2013, 'South Africa's education crisis: The quality of education in South Africa 1994–2011', *Johannesburg: Centre for Development and Enterprise* 21(1), 1–65, viewed 12 March 2021, from https://section27.org.za/wp-content/uploads/2013/10/Spaull-2013-CDE-report-South-Africas-Education-Crisis.pdf.

Spaull, N., Pretorius, E. & Mohohlwane, N., 2020, 'Investigating the comprehension iceberg: Developing empirical benchmarks for early-grade reading in agglutinating African languages', *South African Journal of Childhood Education* 10(1), e1–e14. https://doi.org/10.4102/sajce.v10i1.773

Stahl, S.A. & Murray, B.A., 1994, 'Defining phonological awareness and its relationship to early reading', *Journal of Educational Psychology* 86(2), 221.

Templeton, S., 2004, 'Instructional approached to spelling', in E.R. Silliman & L.C. Wilkinson (eds.), *Language and literacy learning in schools*, pp. 273–291, The Guilford Press, New York, NY.

Theobald, E., 2018, 'Students are rarely independent: When, why, and how to use random effects in discipline-based education research', *CBE Life Sciences Education* 17(3), 1–12. https://doi.org/10.1187/cbe.17-12-0280

Treiman, R., 1991, 'Children's spelling errors on syllable-initial consonant clusters', *Journal of Educational Psychology* 83(3), 346–360.

Van Buuren, S. & Groothuis-Oudshoorn, K., 2011, '{mice}: Multivariate imputation by chained equations in R', *Journal of Statistical Software* 45(3), 1–67, viewed 10 July 2021, from https://www.jstatsoft.org/v45/i03/.

Van der Berg, S., Spaull, N., Wills, G., Gustafsson, M. & Kotzé, J., 2016, *Identifying binding constraints in education*, viewed 10 May 2021, from SSRN 2906945. https://purpletod.co.za/docs/Report%20SA%20Education.pdf

Veber Nielsen, A.-M. & Juul, H., 2016, 'Predictors of early versus later spelling development in Danish', *Reading and Writing: An Interdisciplinary Journal* 29(2), 245–266. https://doi.org/10.1007/s11145-015-9591-y

Veii, K., 2003, *Cognitive and linguistic predictors of literacy in Namibian Herero-English bilingual school children*, University of Surrey, viewed 15 May 2021, from https://core.ac.uk/download/pdf/40107776.pdf.

Weiser, B. & Mathes, P., 2011, 'Using encoding instruction to improve the reading and spelling performances of elementary students at risk for literacy difficulties: A best-evidence synthesis', *Review of Educational Research* 81(2), 170–200. https://doi.org/10.3102/0034654310396719

Wilsenach, C., 2013, 'Phonological skills as predictor of reading success: An investigation of emergent bilingual Northern Sotho/English learners', *Per Linguam: A Journal of Language Learning= Per Linguam: Tydskrif vir Taalaanleer* 29(2), 17–32, https://hdl.handle.net/10520/EJC146853.

Wilsenach, C., 2015, 'Receptive vocabulary and early literacy skills in emergent bilingual Northern Sotho-English children', *Reading & Writing* 6(1), 11. https://doi.org/10.4102/rw.v6i1.77

Wilsenach, C., 2019, 'Phonological awareness and reading in Northern Sotho – Understanding the contribution of phonemes and syllables in grade 3 reading attainment', *South African Journal of Childhood Education* 9(1), e1–e10. https://doi.org/10.4102/sajce.v9i1.647

Woods, A.D., Davis-Kean, P., Halvorson, M.A., King, K.M., Logan, J.A.R., Xu, M. et al., 2021, *Best practices for addressing missing data through multiple imputation*. https://doi.org/10.31234/osf.io/uaezh

Zarić, J., Hasselhorn, M. & Nagler, T., 2020, 'Orthographic knowledge predicts reading and spelling skills over and above general intelligence and phonological awareness', *European Journal of Psychology of Education* 36(1), 21–43. https://doi.org/10.1007/s10212-020-00464-7

Appendix starts on the next page→

Appendix

TABLE A1: Number and proportion of participants per oral reading fluency category in study 1 and study 2.

Category	Study 1: End of Grade 3 N = 49		Study 2: Start of Grade 3 N = 200		Study 2: End of Grade 3 N = 200	
	n	%	*n*	%	*n*	%
0 words correct per minute	4	8.2	60	30.5	14	7
1– 19 words correct per minute	27	55.1	82	41	47	23.5
20– 34 words correct per minute	17	34.7	41	20.5	57	28.5
35+ words correct per minute	1	2	16	8	82	41

Note: Oral reading fluency categories from Ardington et al. (2021) used. According to Ardington et al. (2021) children who read below 20 words correctly per minute are unlikely to be able to read with understanding.