

# Decoding and vocabulary improvements mediate sustained gains in reading comprehension: Evidence from a randomised controlled trial of a multicomponent reading intervention

Cameron Downing,<sup>1,2</sup> Catherine Clark,<sup>3</sup> Gwennant Evans,<sup>3</sup> Rachel Cartin,<sup>3</sup> Joseph Smith,<sup>3</sup> Charles Hulme,<sup>4,5</sup> and Manon Jones<sup>2,3</sup>

<sup>1</sup>Department of Education, University of York, York, UK; <sup>2</sup>Miles Dyslexia Centre, Bangor University, Bangor, UK; <sup>3</sup>Department of Psychology, Bangor University, Bangor, UK; <sup>4</sup>Department of Education, University of Oxford, Oxford, UK; <sup>5</sup>School of Psychology, Oxford Brookes University, Oxford, UK

**Background:** Reading comprehension is critical for academic success, yet many children with persistent decoding difficulties struggle to achieve it. This study examined whether a multicomponent literacy intervention is effective in improving reading comprehension and whether any gains in comprehension are mediated by improvements in word reading and vocabulary knowledge. **Methods:** In a randomised controlled trial (RCT), 285 English-speaking children aged 7–9 years with reading difficulties were assigned to a waitlist control group or the Research Informed Literacy with Language (RILL) intervention, a structured, multicomponent programme targeting decoding and language skills. Literacy outcomes were assessed at baseline (t1), postintervention (t2) and at 4-month follow-up (t3). The trial was preregistered; <https://www.isrctn.com/ISRCTN18940975>. **Results:** Children receiving RILL showed significantly greater gains in word-level literacy ( $d = 0.19$ ,  $p < .001$ ), taught vocabulary ( $d = 0.30$ ,  $p = .017$ ) and reading comprehension ( $d = 0.23$ ,  $p = .011$ ) immediately postintervention. Effects were sustained at follow-up (word-level literacy  $d = 0.17$ ; taught vocabulary  $d = 0.30$ ; comprehension  $d = 0.25$ ). Mediation analyses, showed a significant indirect effect of the intervention on comprehension at delayed follow-up via word-level literacy at t2 (y-standardised indirect  $\beta = .10$ , 95% CI [0.06, 0.12]), with a negligible direct effect (y-standardised  $\beta = .01$ , 95% CI [-0.20, 0.20]). In an additional exploratory parallel-mediation model, both t2 word-level literacy and taught vocabulary showed unique indirect effects on t3 comprehension (word-level literacy indirect:  $\beta = .16$ , 95% CI 0.10, 0.21, and taught vocabulary indirect  $\beta = .11$ , 95% CI 0.02, 0.23). **Conclusions:** Our intervention produced immediate and sustained improvements in word-level literacy, taught vocabulary and reading comprehension in struggling readers. Persisting decoding weaknesses are common in later primary years, and our findings show that improving word reading can produce enduring benefits for comprehension. **Keywords:** Reading comprehension; decoding; intervention; longitudinal mediation; struggling readers; literacy development.

Reading comprehension is a foundational academic skill that underpins learning across subject areas and predicts long-term educational success (Snow, 2002). The Simple View of Reading (SVR; Gough & Tunmer, 1986) and the Reading is Language model (RiL; Snowling & Hulme, 2025) provide influential frameworks for understanding how comprehension develops. Both models propose that reading comprehension reflects the combined contributions of word decoding and oral language comprehension. A substantial body of evidence supports the predictive roles of both components across languages, including English (Catts, Adlof, & Weismer, 2006; Keenan, Betjemann, & Olson, 2008; Salceda, Alonso, & Castilla-Earls, 2014) and more transparent orthographies such as Dutch, Finnish, Greek and Norwegian (Caravolas et al., 2019; de Jong & van der Leij, 2002; Florit & Cain, 2011; Kendeu, Papadopoulos, & Kotzapoulou, 2013; Lervåg, Hulme, & Melby-Lervåg, 2018; Müller & Brady, 2001).

Cross-linguistic research highlights the role of orthographic depth in shaping this balance. In opaque orthographies such as English, decoding is relatively difficult to master and strongly predicts comprehension in the early years, whereas in transparent orthographies decoding is mastered earlier, allowing language comprehension to play a larger role from a younger age (Caravolas, Lervåg, Defior, Seidlová Málková, & Hulme, 2013; Caravolas et al., 2019; Florit & Cain, 2011; García & Cain, 2014; Leppänen et al., 2004; Verhoeven, van Leeuwe, & Vermeer, 2011; Wimmer & Goswami, 1994). Even in English, however, this balance shifts developmentally: decoding constrains comprehension early on, but as fluency develops, language comprehension becomes the primary driver (Adlof, Catts, & Lee, 2010; Kim, Wagner, & Foster, 2011; Tilstra, McMaster, Van den Broek, Kendeou, & Rapp, 2009).

For struggling readers, this trajectory is disrupted. Persistent decoding difficulties limit fluency and restrict the cognitive resources available for higher-order comprehension processes (Leach, Scarborough, & Rescorla, 2003; Ouellette & Beers, 2010).

Conflict of interest statement: No conflicts declared.

Longitudinal research shows that decoding difficulties often persist into later grades, constraining comprehension longer than for typically developing peers (Catts, 2003; Torppa et al., 2007). Decoding also remains a significant predictor of comprehension growth in poor readers after it ceases to be predictive in more skilled readers (Lervåg & Hulme, 2010). Lervåg et al. (2018) further demonstrated a curvilinear relationship, with decoding exerting the strongest effects for children with the lowest skills and weakening once sufficient accuracy and fluency are achieved.

These findings point to decoding as a developmental bottleneck for poorer readers and raise the question of whether intervention-induced gains in decoding can unlock later gains in comprehension, as predicted by the SVR and RiL models. While phonics-based interventions reliably improve word reading (Ehri et al., 2001; Hatcher, Hulme, & Ellis, 1994; Torgesen et al., 2001), effects on reading comprehension are mixed and often modest or delayed (Clarke, Snowling, Truelove, & Hulme, 2010; Fricke, Bowyer-Crane, Haley, Hulme, & Snowling, 2013). This has led to calls for multicomponent programmes targeting both decoding and comprehension (Silverman, Johnson, Keane, & Khanna, 2020). Meta-analyses indicate that decoding instruction exerts the strongest influence on comprehension outcomes, even after accounting for other skills (Melby-Lervåg & Lervåg, 2014), and recent trials identify decoding as a key mediator of programme effects (Lovett et al., 2022; Vousden et al., 2022), though effects are attenuated in older children.

Despite this evidence, most findings remain correlational. Longitudinal studies show that early decoding predicts later comprehension (e.g. Adlof et al., 2010; Catts et al., 2006; Lervåg, 2021; Lervåg et al., 2018), but cannot establish causality. Mediation analyses embedded in intervention studies offer a stronger test of mechanisms of change (Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012), yet few have used temporally lagged designs. Ahmed, Miciak, Taylor, and Francis (2022) examined changes in relations among component skills but measured mediators only at baseline. Fricke et al. (2013) demonstrated a lagged indirect effect via language gains, whereas Vousden et al. (2022) approximated lagged mediation using earlier literacy gains to predict later national outcomes. Roberts, Vaughn, Fletcher, Stuebing, and Barth (2013) modelled growth but did not test mediation. Temporally lagged mediation within an RCT therefore provides a particularly stringent test of whether decoding gains drive later comprehension growth.

Embedding such designs also provides a strong test of theory. The RiL model (Snowling & Hulme, 2025) proposes that later comprehension depends on earlier word reading and language skills, yet to our knowledge, no intervention study has embedded a temporally lagged mediation design to

test whether immediate postintervention decoding gains lead to later comprehension improvements.

### The current study

We conducted a randomised controlled trial (RCT) evaluating the Research Informed Literacy with Language (RILL) intervention in struggling readers aged 7–9 years. RILL combines structured decoding instruction with comprehension and language activities. We asked whether RILL produced immediate (t2) and sustained (t3) gains in decoding, vocabulary and comprehension, and whether later comprehension gains were mediated by earlier improvements in decoding.

## Method

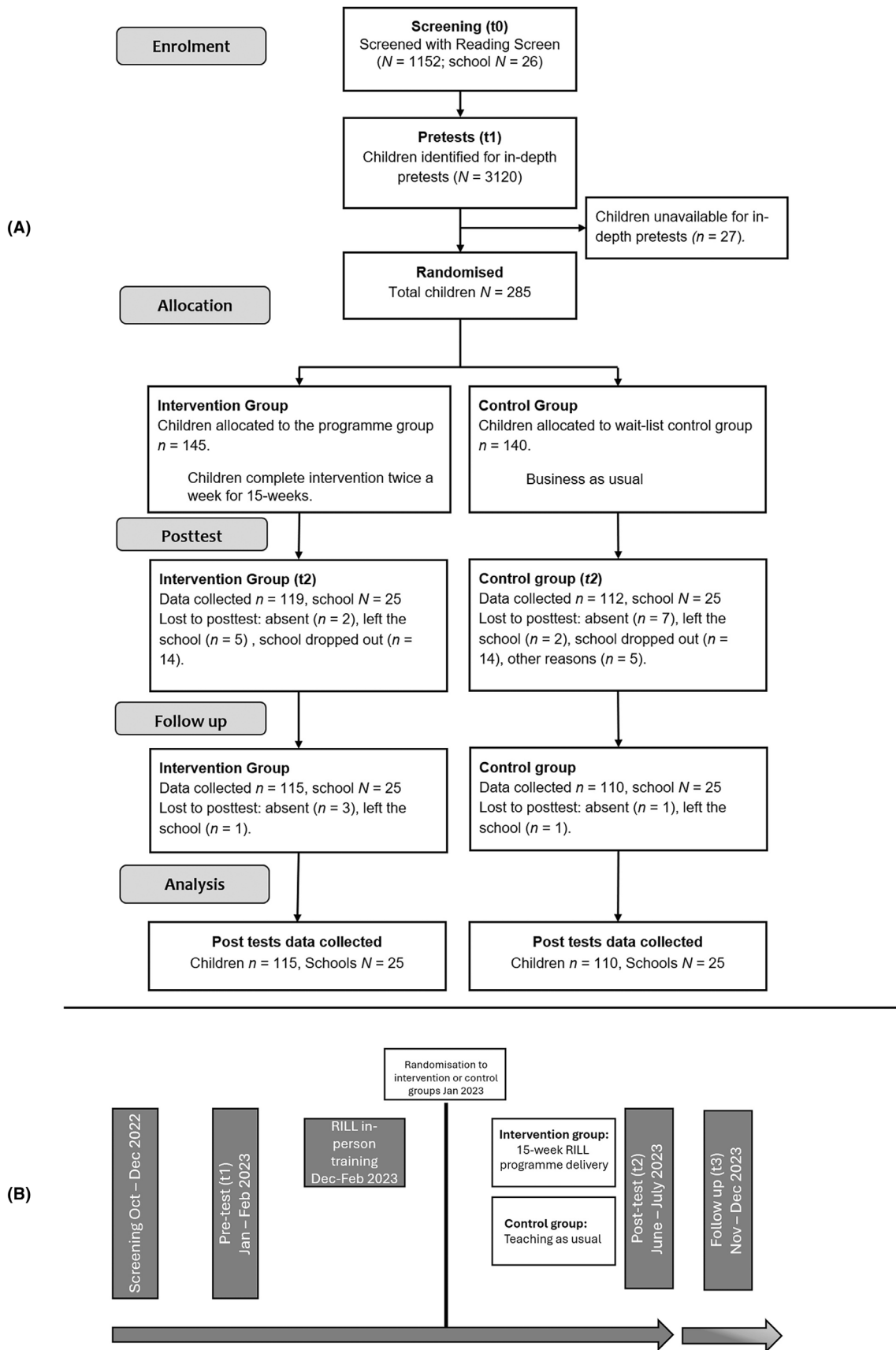
### Design and participants

Twenty-six schools were recruited to participate in the study. In these UK schools 1,152 children in Years 3 and 4 were screened for eligibility using ReadingScreen, an app-based assessment of word and nonword reading (<https://oxedandassessment.com/readingscreen/>). The 12 lowest-performing readers in each school (six in each class) were randomly assigned within classrooms to the intervention (RILL;  $n = 145$ ; mean age = 100.82 months;  $SD = 9.98$ ) or a wait-list control condition ( $n = 140$ , mean age = 101.11 months;  $SD = 7.57$ ). Children were selected as the lowest-performing readers within each class and as a result, at baseline (t1), participants' word reading was substantially below age expectations. On the WRAT-5 Word Reading subtest (normed mean = 100,  $SD = 15$ ), the sample had a mean standard score of  $M = 64.20$  ( $SD = 9.67$ , range = 55–97). Eighty-two per cent of the sample scored  $\leq 85$  ( $\geq 1$   $SD$  below the normative mean), and 73% scored  $\leq 70$  ( $\geq 2$   $SD$ s below the normative mean), indicating that a majority of participants were in the low-achievement range.

Randomisation was carried out by an independent statistician. Headteachers provided consent for school participation, and school administrators distributed invitation letters to caregivers to request consent for their child to participate. Pretest assessments (t1) of reading and related language skills were administered via the platform Gorilla (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020). Posttest assessments (t2) were completed 15 weeks later after the intervention group had completed the programme, and the follow-up assessment (t3) was conducted 4 months later. Participant flow is shown in the CONSORT diagram (Figure 1a; Schulz, Altman, & Moher, 2010). Figure 1b provides a timeline for the project. See also our preregistration of the trial here: <https://www.isrctn.com/ISRCTN18940975>. Testing was done on site during school hours. During data collection, we experienced technical issues in collecting planned language measures; this led to us dropping a hypothesis about the effects of the intervention on language. However, we were still able to collect expressive vocabulary measures, and so we include analysis of the effects of the intervention on this measure as an exploratory analysis. No other changes to methods were made following trial registration, and the trial end was as planned.

### The RILL intervention

The RILL intervention (Research Informed Literacy with Language; Downing et al., 2025) is a 15-week, twice-weekly,



**Figure 1** Panel A. CONSORT diagram showing the flow of participants through the RCT. Panel B. Timeline of the trial indicating the times of assessment, training and intervention

individualised programme (totalling 20 h). It builds on previously validated interventions, including REVI (Duff et al., 2008) and Reading Intervention (Bowyer-Crane et al., 2008; Hatcher et al., 1994; Hatcher, Hulme, & Snowling, 2004), with a focus on phonological awareness, vocabulary and decoding. Lessons, delivered online by trained teaching assistants (TAs), incorporate distributed learning principles (Latimier, Peyre, & Ramus, 2021).

RILL was designed to strengthen core literacy skills in children aged 7–9 years with persistent reading difficulties, particularly vocabulary and word-level decoding skills that underpin comprehension (Scarborough, 2001). Teaching Assistants (TAs) delivered the RILL lessons twice a week for 15 weeks in small groups in school (2–3 children at a time). The TAs led the small group through interactive activities presented on a computer.

Each 40-min lesson follows a structured sequence:

- 1 Words of the Day (5 min): Children are taught the meanings of 2 Tier 2 vocabulary words using definitions, visual support and discussion (Beck et al., 1982, 2002).
- 2 Passage of the Day (10 min): Children read aloud a short passage containing the target words.
- 3 Word Games (5 min): Activities to support phonemic awareness and phonics (e.g. blending).
- 4 Word Writing (10 min): Spelling instruction with a focus on vowel patterns and sound-letter correspondence.
- 5 Story Time (5 min): Narrative skills practice using scaffolded story creation tasks (Clarke et al., 2010).
- 6 Recap (5 min): Children reviewed and applied the target vocabulary.

Schools continued to provide business-as-usual instruction and support to pupils allocated to the wait-list control arm and any additional literacy support that pupils were already receiving continued.

### Ethical information

Written informed consent was obtained from parents/caregivers for all participating children, with child assent also sought in accordance with ethical guidelines. The study received ethical approval from the School of Psychology ethics board at Bangor University, granted on 07.09.2022, reference number 2022-17213.

### Assessment measures

**Selection measure.** The screening tool used to select children for the trial was *ReadingScreen* (<https://oxedandassessment.com/readingscreen/>). This mobile app containing two subtests: word and nonword reading. The child is asked to read aloud a series of 60 words and 34 nonwords of increasing difficulty. The assessment took less than 10 min to complete per child and was administered, following training, by the class teacher or TA.

**Primary outcomes.** The primary outcome measures are two latent literacy variables reflecting: (1) Word-level literacy (created from measures of single word reading, single word spelling, word fluency and pseudoword fluency) and (2) Reading Comprehension.

**Word-level literacy measures.** Word Reading Accuracy (*Wide Range Achievement Test 5<sup>th</sup> edition* WRAT-5; Wilkinson & Robertson, 2017): This untimed test measures letter identification and word recognition. The child reads each item aloud. The words are of increasing complexity in terms of regularity and word frequency, with a total of 15 letters and 55 words. The task is discontinued following 10 consecutive errors.

Word Reading Accuracy (ReadingScreen) <https://oxedandassessment.com/readingscreen/>: This untimed test (described above) measures word recognition accuracy. The child reads each item aloud. The words are of increasing complexity in terms of regularity and word frequency, with 60 items in total. Responses are marked as correct or incorrect by the administrator.

Word Spelling Accuracy (WRAT-5): This test measures the child's ability to write letters and words to dictation without a time limit. Words increase in complexity in terms of spelling patterns with a total of 13 letters and 42 words. The task is discontinued following 10 consecutive errors.

Passage Reading Accuracy (*York Assessment of Reading for Comprehension Passage Reading; YARC Passage Reading*; Snowling et al., 2009): To measure word reading accuracy during passage reading, children read aloud four passages (Levels) of the YARC: at t1, t2 and t3. One point is awarded for each correct word. At each time point, the accuracy scores from each passage were used to produce a passage reading accuracy score. Administering the same four passages to all participants at each time point was designed to capture the full range of ability (c.f., Nation, Cocksey, Taylor, & Bishop, 2010).

**Reading comprehension measure.** YARC Comprehension Questions: (*YARC Passage Reading*; Snowling et al., 2009): To measure reading comprehension, after reading each of the passages aloud (see Passage reading accuracy), children are asked a series of comprehension questions based on the passage. Published scoring guidelines were followed when scoring responses to each of the comprehension questions.

**Secondary outcome.** A short expressive vocabulary task was used to assess knowledge of the vocabulary explicitly taught over the course of the programme. A subset of 18 words drawn from the 60 words taught during the RILL lessons was presented to the child aurally. The child was then required to briefly define the meaning of each word. Children's definitions were scored on a 3 point scale (0—*incorrect, or no response*; 1—*partially correct definition*; 2—*fully correct definition*).

### Procedure

Following written caregiver consent, children completed baseline testing (t1; ~1 h). Teaching Assistants, trained by the research team, administered identical assessments at all time points in school during school hours. Assessments were administered using the Gorilla platform (<https://gorilla.sc>), with support from research assistants. Posttesting occurred after 15 weeks (t2) and again 4 months later (t3). Assessors were blind to group assignment at t2 and t3, and 15% of all data was rescored to ensure interrater reliability. Intervention began in early 2023.

Implementation fidelity was assessed via a single, structured 45-min observation per school by a trained Research Assistant. Each lesson component (words of the day, passage of the day, word games, spelling, story writing, revisiting words of the day), as well as overall lesson timing, pupil engagement and use of technology, was rated on a 3-point scale (1 = *minimal*, 2 = *partial*, 3 = *full implementation*). Observers also recorded qualitative notes, and Teaching Assistants were invited to provide feedback. All schools achieved total scores >20 (out of 27), indicating high fidelity.

### Trial registration

This trial was preregistered with the ISRCTN registry (ISRCTN18940975; registered 20/06/2023; <https://www.isrctn.com/ISRCTN18940975>).

## Results

Analyses were conducted in Stata® 18.5 (Stata Corporation, College Station, TX, USA, 2023) and structural equation models (SEM) were implemented in Mplus 8.11 using full information maximum likelihood (FIML) estimators to handle missing data. All analyses were performed on an intention-to-treat basis. The study was preregistered (<https://www.isrctn.com/ISRCTN18940975>) and the data and analysis files are available at: [https://osf.io/5vmqj/?view\\_only=92f4538009ca4657ab08975d03b364de](https://osf.io/5vmqj/?view_only=92f4538009ca4657ab08975d03b364de).

A total of 1,152 children in Years 3 and 4 were screened with ReadingScreen at t0 across 26 schools. Children were randomly allocated within the school to the intervention or control group (6 children per arm). At posttest (t2), 26 children from the intervention group and 28 children from the control group were lost. At follow-up (t3), a further 4 children from the intervention group and 2 children from the control group were lost. There were no significant differences at pretest between the children who completed the study and those who were lost in age ( $p = .391$ ,  $d = .13$ ) or ReadingScreen scores ( $p = .937$ ,  $d = .01$ ). Table 1 shows the descriptive statistics for all measures at screening (t0), pretest (t1), posttest (t2) and follow-up (t3) for both groups. Both groups performed similarly at pretest as expected given random assignment.

### Primary outcome effects

Our primary outcomes were word-level literacy and reading comprehension. Intervention effects were estimated using ANCOVA models in which the relevant latent outcome at posttest (t2; short-term effects) or follow-up (t3; longer-term effects) was regressed on its pretest value (t1) and a dummy-coded group variable (intervention = 1, control = 0). Prior to these analyses, confirmatory factor analyses (CFAs) were conducted to evaluate model fit and test factorial invariance across time. Residuals for the same indicators were correlated across time points, which improved the model fit. Clustering within schools was accounted for using robust (Huber-White) standard errors with school as the clustering variable.

For each model, we tested the homogeneity-of-slopes assumption by comparing models with group-specific versus constrained pretest-outcome slopes. In all cases, the constrained model was supported, indicating no evidence of a group-by-pretest interaction; intervention effects did not vary as a function of baseline ability. Intervention effects are reported as  $y$ -standardised regression coefficients for the group variable, equivalent to Cohen's  $d$ .

**Immediate effects on word literacy.** The word-level literacy latent variable was defined using four measures (WRAT Word Reading, WRAT Word

Spelling, ReadingScreen Word Reading Fluency and YARC Reading Accuracy). Residuals were correlated across time for all indicators except WRAT Word Reading. Tests of longitudinal invariance supported full configural, metric and scalar invariance (see Appendix S1, Tables S1 and S2), and the scalar model was therefore used. The ANCOVA model (Figure 2, Panel A) showed good fit,  $\chi^2(31) = 45.85$ ,  $p = .042$ , RMSEA = 0.041 (90% CI [0.01, 0.07]), SRMR = 0.082, CFI = 0.99, TLI = 0.99. Children in the intervention group made significantly greater gains in word-level literacy immediately postintervention than controls ( $d = 0.19$ , 95% CI [0.09, 0.26],  $p < .001$ ).

**Longer-term effects on word literacy.** Longer-term effects were examined using word-level literacy latent variables derived from the four word reading and spelling measures at pretest (t1) and follow-up (t3). Residuals were correlated for WRAT Word Spelling and YARC Reading Accuracy across time. Tests supported full configural, metric and scalar invariance (see Appendix S1, Tables S1 and S2), and the scalar model was therefore used (Figure 2, Panel B). Model fit was good,  $\chi^2(31) = 42.19$ ,  $p = .087$ , RMSEA = 0.036 (90% CI [0.00, 0.06]), SRMR = 0.094, CFI = 0.99, TLI = 0.99. Children in the intervention group showed significantly greater gains in word-level literacy than controls at follow-up ( $d = 0.17$ , 95% CI [0.05, 0.28],  $p = .004$ ), an effect comparable in size to the immediate postintervention effect, indicating maintenance of gains.

**Immediate effects on reading comprehension.** The reading comprehension latent variable was derived from comprehension question scores for four YARC passages. Residuals for Passages 3 and 4 were correlated across time, and passages were correlated within the pretest to account for shared task variance and improve model fit. Full configural and partial metric invariance (three noninvariant loadings) were supported (see Appendix S2, Tables S3 and S4). The partial metric model was used to estimate intervention effects. The ANCOVA model (Figure 3, Panel A) was a good fit to the data,  $\chi^2(23) = 35.62$ ,  $p = .045$ , RMSEA = 0.045 (90% CI [0.01, 0.07]), SRMR = 0.087, CFI = 0.98, TLI = 0.98. Children in the intervention group showed significantly greater gains in reading comprehension at posttest than controls ( $d = 0.23$ , 95% CI [0.06, 0.42],  $p = .011$ ).

**Longer-term effects on reading comprehension.** Longer-term intervention effects on reading comprehension were tested using latent variables derived from four YARC passages administered at pretest (T1) and follow-up (T3). Residuals for Passages 2–4 were correlated across time, and passages were correlated within pretest to account for shared task

**Table 1** Mean, standard deviations and reliabilities of raw scores for the intervention and control groups for the outcome measures at screening (t0), pretest (t1), posttest (t2) and follow-up (t3)

	Intervention			Control Group			Reliability
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	
Age							
t0	145	100.82	9.98	140	101.11	7.57	
Word literacy							
Reading screen							.95 <sup>a</sup>
t0	116	16.67	10.26	111	17.72	9.98	
t2	68	24.46	10.74	68	24.73	11.48	
t3	76	26.37	11.64	76	27.12	10.49	
WRAT reading							.95–.96 <sup>a</sup>
t1	124	9.48	5.67	122	10.66	5.69	
t2	115	11.59	5.95	107	11.78	6.23	
t3	113	13.64	5.75	106	13.84	6.28	
WRAT spelling							.91–.92 <sup>a</sup>
t1	143	5.13	2.80	137	5.18	2.58	
t2	115	6.29	3.52	112	6.03	3.12	
t3	113	7.12	3.34	109	6.82	3.29	
YARC reading accuracy							.85–.93 <sup>b</sup>
t1	134	111.44	67.07	124	125.77	67.79	
t2	111	208.29	96.56	102	201.09	97.12	
t3	110	225.09	87.80	105	222.87	86.20	
Reading comprehension							
Passage 1							.55–.69 <sup>b</sup>
t1	118	3.03	1.92	115	3.63	1.76	
t2	89	3.93	2.05	89	4.01	2.18	
t3	95	4.13	2.00	95	4.18	2.29	
Passage 2							
t1	118	3.50	2.06	117	4.13	2.31	
t2	95	2.84	2.14	85	2.51	1.92	
t3	96	3.18	1.98	93	3.19	2.24	
Passage 3							
t1	118	2.90	1.73	102	3.38	1.67	
t2	107	3.45	1.71	93	3.66	1.65	
t3	106	3.75	1.71	97	3.77	1.62	
Passage 4							
t1	113	1.46	1.66	100	1.89	1.68	
t2	101	2.24	1.90	86	2.45	1.75	
t3	102	2.84	2.11	97	2.94	1.75	
Language							
Vocabulary 1							.77–.85 <sup>b</sup>
t1	119	2.30	1.53	115	2.10	1.72	
t2	105	3.37	2.04	102	2.31	1.97	
t3	110	3.58	2.14	109	2.80	2.20	
Vocabulary 2							
t1	119	1.22	1.55	115	1.26	1.55	
t2	105	2.62	2.28	102	2.02	2.31	
t3	110	2.82	2.35	109	1.93	2.14	
Vocabulary 3							
t1	115	3.47	2.02	115	3.22	2.33	
t2	102	4.66	2.31	102	3.74	2.60	
t3	109	4.88	2.40	109	3.95	2.65	

Reliabilities are reported as Cronbach's alpha.

<sup>a</sup>Published reliabilities.

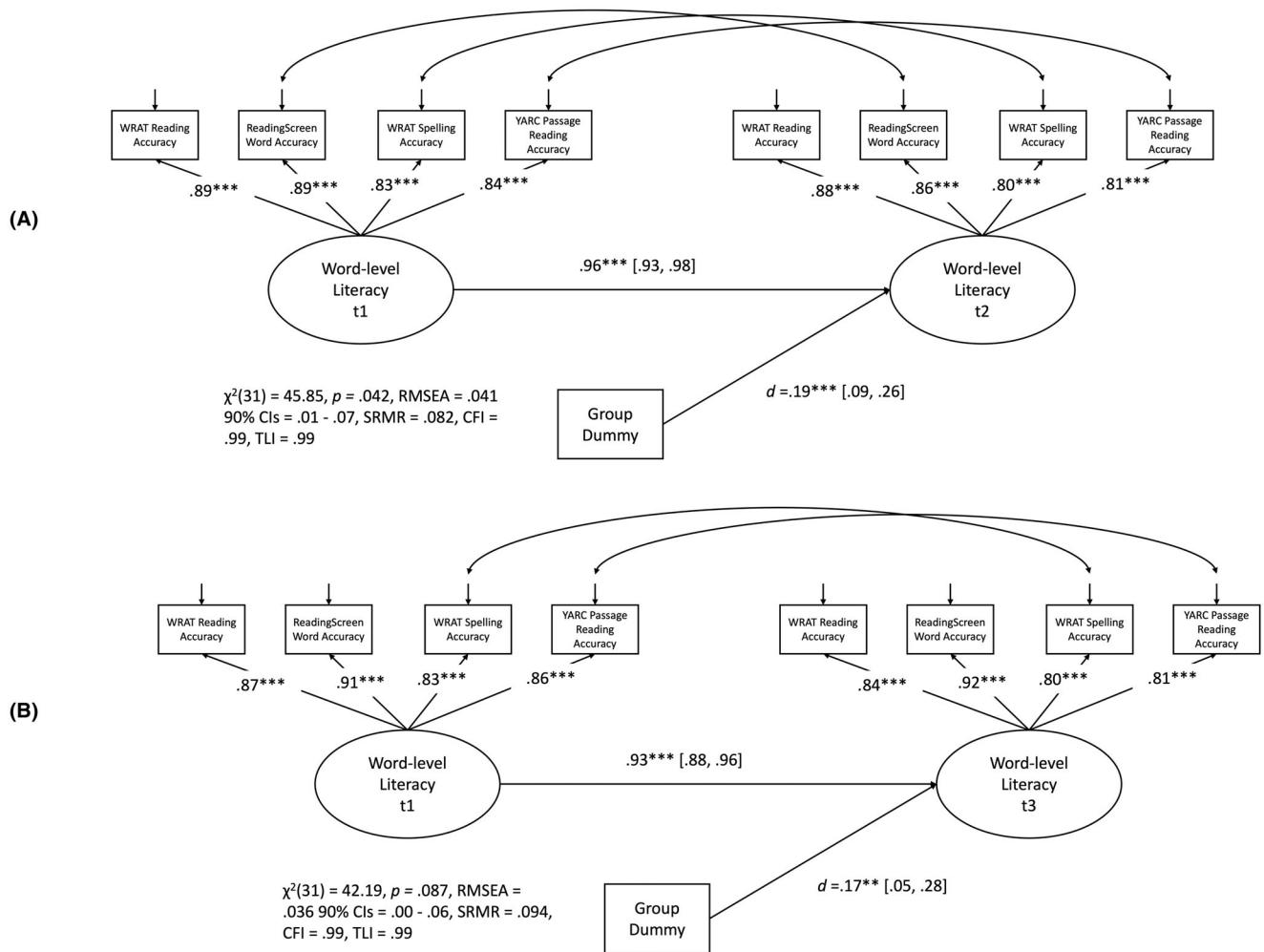
<sup>b</sup>Sample reliabilities.

variance. Full configural and partial metric invariance (three noninvariant loadings) were supported (see Appendix S2, Tables S3 and S4), and the partial metric model was therefore used. The ANCOVA model (Figure 3, Panel B) showed an adequate fit,  $\chi^2(22) = 42.04$ ,  $p = .006$ , RMSEA = 0.057 (90% CI [0.03, 0.08]), SRMR = 0.088, CFI = 0.97, TLI = 0.95. Children in the intervention group showed significantly greater reading comprehension gains at

follow-up than controls ( $d = 0.25$ , 95% CI [0.03, 0.47],  $p = .026$ ).

### Secondary outcome effects

As with the primary outcome measures, we measured the effects of the intervention using the ANCOVA approach described earlier. This analysis was not preregistered.



**Figure 2** Panel A. Path diagram for latent variables of word-level literacy at pretest (t1) and posttest (t2). Factor loadings and the autoregressive path are standardised estimates. The effect of the intervention is shown by the path from the group dummy (intervention vs. control) to word-level literacy at posttest, which is y-standardised (equivalent to Cohen's  $d$ ). CIs are 95% robust 95% CIs accounting for clustering within schools are reported in square brackets. \*\*\* $p < .01$ . Panel B. Path diagram for latent variables of word-level literacy at pretest (t1) and follow-up (t3). Factor loadings and the autoregressive path are standardised estimates. The effect of the intervention is shown by the path from the group dummy (intervention vs. control) to word-level literacy at follow-up, which is y-standardised (equivalent to Cohen's  $d$ ). CIs are 95% robust 95% CIs accounting for clustering within schools are reported in square brackets. \*\* $p < .01$ , \*\*\* $p < .01$ .

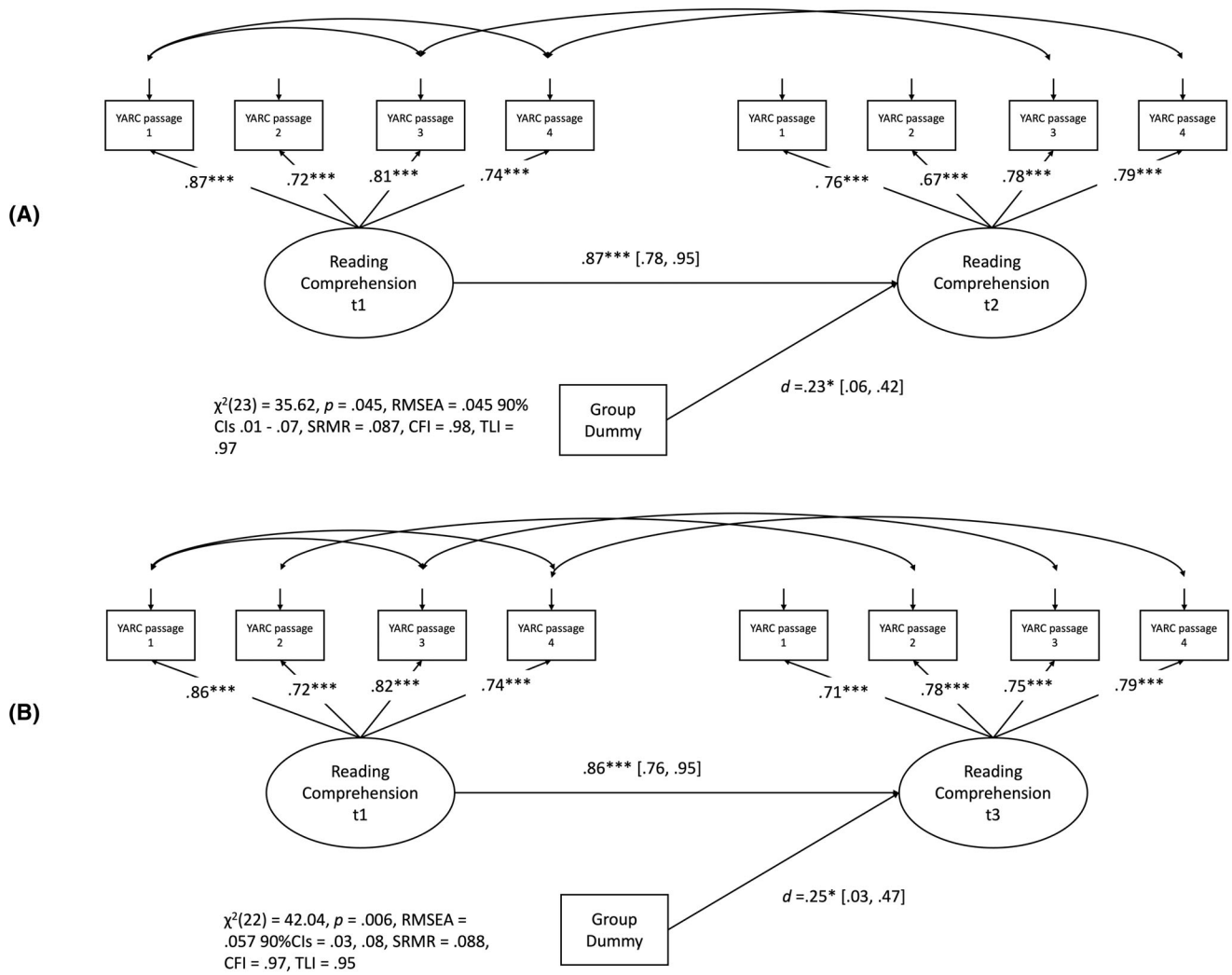
**Immediate effects on vocabulary.** The expressive vocabulary latent variable was defined using three parcels (six words per parcel) at each time point. Residuals were correlated across time for all parcels except Parcel 1. Full configural, metric and scalar invariance was supported (see Appendix S3, Tables S5 and S6), and the scalar model was used to estimate intervention effects. The ANCOVA model (Figure 4, Panel A) showed good fit,  $\chi^2(17) = 25.28, p = .089, RMSEA = 0.043$  (90% CI [0.00, 0.08]), SRMR = 0.088, CFI = 0.98, TLI = 0.98. Children in the intervention group made significantly greater gains in taught vocabulary than controls at posttest ( $d = 0.30, 95\% \text{ CI } [0.05, 0.54], p = .017$ ).

**Longer-term effects on vocabulary.** Longer-term effects on expressive vocabulary were examined using latent variables at pretest (T1) and follow-up (T3). Residuals for Parcel 2 were correlated across time. Full configural, metric and scalar invariance

was supported (see Appendix S3, Tables S5 and S6), and the scalar model was used to test group effects. The model (Figure 4, Panel B) showed good fit,  $\chi^2(17) = 30.12, p = .026, RMSEA = 0.053$  (90% CI [0.02, 0.08]), SRMR = 0.093, CFI = 0.98, TLI = 0.97. The intervention group showed significantly greater vocabulary gains than controls at follow-up ( $d = 0.30, 95\% \text{ CI } [0.04, 0.43], p = .021$ ), with effects comparable to those observed at posttest, indicating maintenance.

### Testing mediators of gains in reading comprehension

We tested whether later reading comprehension gains were mediated by earlier improvements in word-level literacy using a temporally lagged mediation model. Word-level literacy at pretest (T1) and posttest (T2) was modelled as a latent construct defined by WRAT Word Reading, WRAT Word Spelling, ReadingScreen word



**Figure 3** Panel A. Path diagram for latent variables of reading comprehension at pretest (t1) and posttest (t2). Factor loadings and the autoregressive path are standardised estimates. The effect of the intervention is shown by the path from the group dummy (intervention vs. control) to word-level literacy at posttest, which is y-standardised (equivalent to Cohen’s *d*). CIs are 95% robust 95% CIs accounting for clustering within schools are reported in square brackets. \*\**p* < .01, \*\*\**p* < .01. Panel B. Path diagram for latent variables of reading comprehension at pretest (t1) and follow-up (t3). Factor loadings and the autoregressive path are standardised estimates. The effect of the intervention is shown by the path from the group dummy (intervention vs. control) to reading comprehension at follow-up, which is y-standardised (equivalent to Cohen’s *d*). CIs are 95% robust 95% CIs accounting for clustering within schools are reported in square brackets. \*\**p* < .01, \*\*\**p* < .01

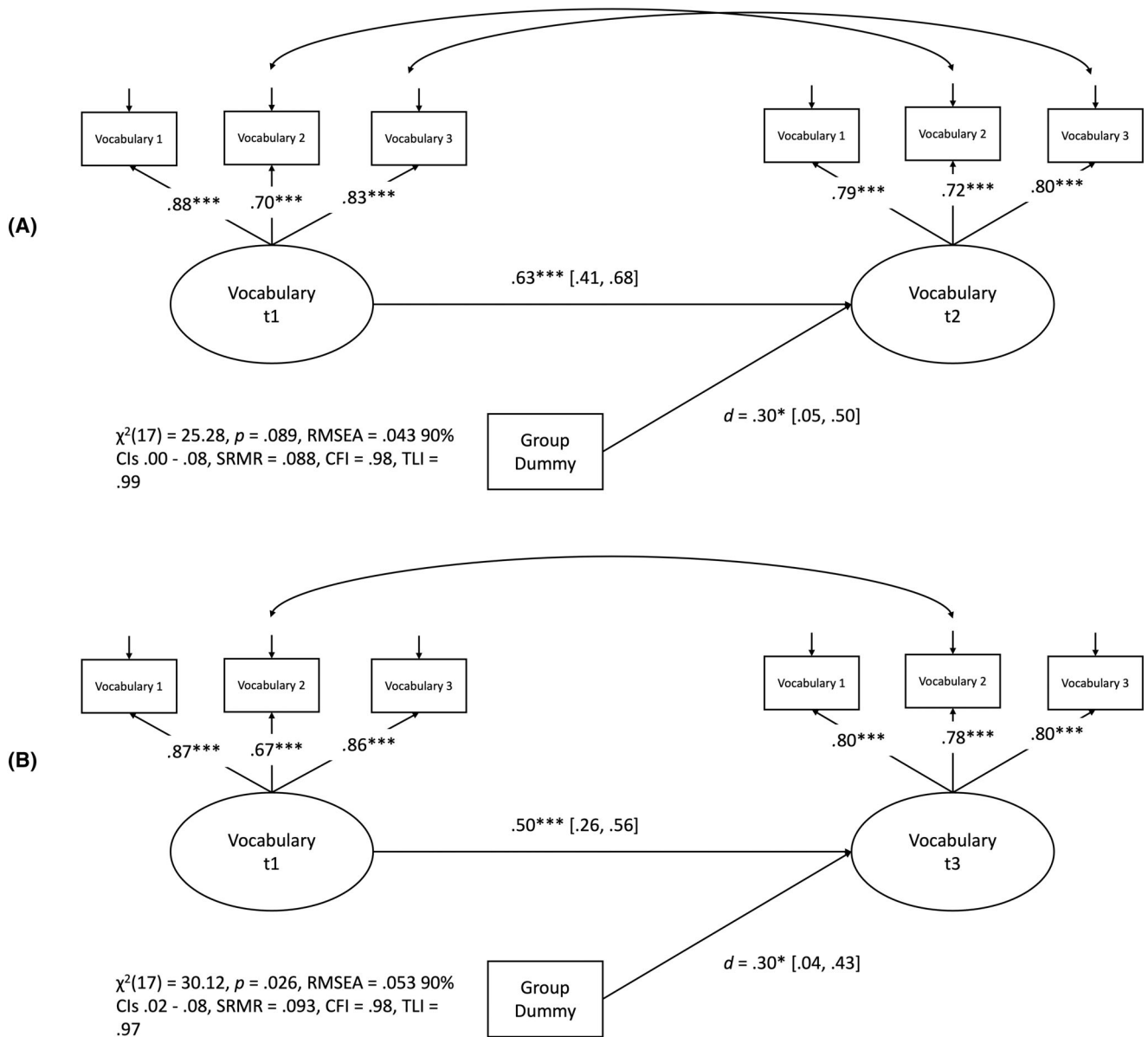
reading and YARC passage reading accuracy, with residuals correlated across time (except WRAT Word Reading). Reading comprehension at follow-up (T3) was defined by the four YARC passage comprehension scores. The model regressed T1 word-level literacy and group onto T2 word-level literacy, and T2 word-level literacy and group onto T3 reading comprehension (Figure 5a).

The model provided a good fit to the data,  $\chi^2(65) = 109.49, p < .001, RMSEA = 0.049$  (90% CI [0.03, 0.07]), SRMR = 0.086, CFI = 0.98, TLI = 0.97. Children in the intervention group showed significantly greater gains in word-level literacy at posttest than controls ( $d = 0.22$ ). The direct effect of the intervention on reading comprehension at follow-up was negligible (y-standardised = 0.01, 95% CI [−0.20, 0.19]); however, the indirect effect via posttest word-level literacy was significant (standardised

indirect effect = 0.10, 95% CI [0.06, 0.12],  $p < .001$ , equivalent to  $d = 0.19$  (Kraft, 2020).

Given the intervention’s multicomponent nature, improvements in reading comprehension could be an additional, exploratory parallel-mediation model including vocabulary at posttest (t2) as a proxy for language skills alongside word-level literacy to test whether each uniquely mediated follow-up reading comprehension (t3; Figure 5b). We note that vocabulary captures only one aspect of oral language and that the study was not specifically powered to compare multiple mediators. These analyses are strictly exploratory.

This model provides an appropriate fit to the data,  $\chi^2(151) = 262.76, p < .001, RMSEA = 0.051$  90% CI = [0.04, 0.06], SRMR = 0.077, CFI = 0.96, TLI = 0.95. In this model, accounting for later reading comprehension, we found that children who completed the intervention made educationally



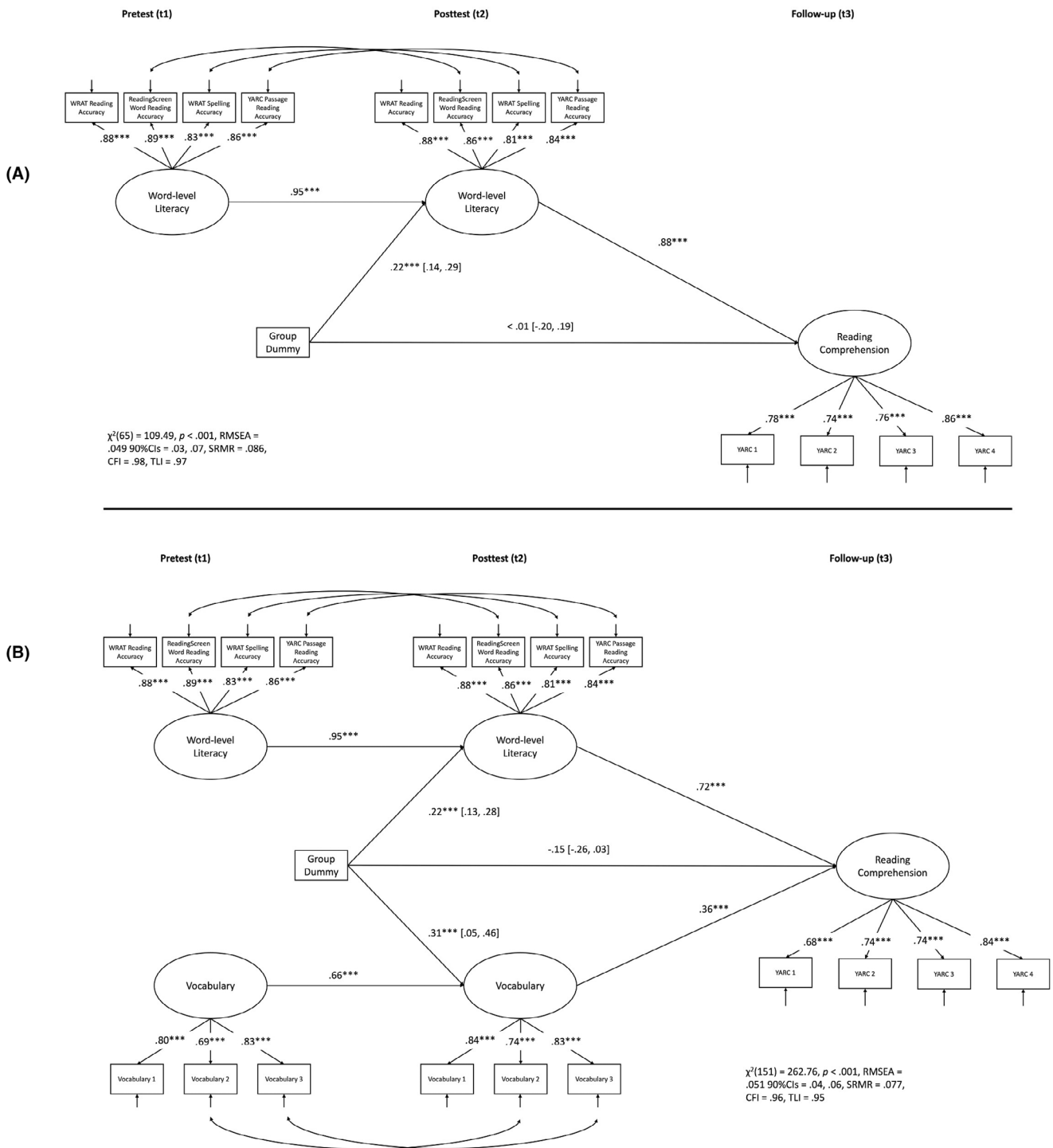
**Figure 4** Panel A. Path diagram for latent variables of taught expressive vocabulary at pretest (t1) and posttest (t2). Factor loadings and the autoregressive path are standardised estimates. The effect of the intervention is shown by the path from the group dummy (intervention vs. control) to word-level literacy at posttest, which is y-standardised (equivalent to Cohen's *d*). CIs are 95% robust CIs accounting for clustering within schools are reported in square brackets.  $**p < .01, ***p < .01$ . Panel B. Path diagram for latent variables of expressive vocabulary at pretest (t1) and follow-up (t3). Factor loadings and the autoregressive path are standardised estimates. The effect of the intervention is shown by the path from the group dummy (intervention vs. control) to reading comprehension at follow-up, which is y-standardised (equivalent to Cohen's *d*). CIs are 95% robust CIs accounting for clustering within schools are reported in square brackets.  $**p < .01, ***p < .01$

significantly greater gains in their word-level literacy ( $d = 0.22$ ) and vocabulary ( $d = 0.31$ ) than children who did not complete the intervention (Kraft, 2020). The direct effect of the intervention on later reading comprehension in this model was nonsignificant (y-standardised effect =  $-0.15$  95% CI  $[-0.34, 0.05]$ ). The indirect effect of the intervention on reading comprehension via posttest word-level literacy was significant (intervention  $\rightarrow$  posttest word-level literacy  $\rightarrow$  follow-up reading comprehension; standardised indirect effect =  $0.16$ ; 95% CI  $[0.10, 0.21]$ ;  $p < .001$ ) as was the indirect effect of the intervention on reading comprehension via posttest

vocabulary (intervention  $\rightarrow$  posttest vocabulary  $\rightarrow$  follow-up reading comprehension; standardised indirect effect =  $0.11$ ; 95% CI  $[0.02, 0.23]$ ;  $p = .048$ ).

## Discussion

The present study evaluated a language-rich reading intervention programme and its effects on both reading accuracy and reading comprehension. A particular focus was whether intervention-induced gains in decoding could account for later improvements in reading comprehension in children with persistent decoding difficulties.



**Figure 5** Panel A. Path diagram of the mediation model testing the indirect influence of the intervention on reading comprehension via word-level literacy. Factor loadings and the autoregressive path are standardised estimates. The effect of the intervention is shown by the paths from the group dummy (intervention vs. control) to word-level literacy and reading comprehension at follow-up, which is y-standardised (equivalent to Cohen’s *d*). Panel B. Path diagram of the mediation model testing the indirect influence of the intervention on reading comprehension via word-level literacy and vocabulary. Factor loadings and the autoregressive path are standardised estimates. The effect of the intervention is shown by the paths from the group dummy (intervention vs. control) to word-level literacy, vocabulary and reading comprehension at follow-up are y-standardised (equivalent to Cohen’s *d*). CIs are 95% robust CIs accounting for clustering within schools are reported in square brackets. \*\**p* < .01, \*\*\**p* < .01

Consistent with prior multicomponent intervention trials, children who received RILL made significantly greater gains than controls in word decoding, taught vocabulary knowledge and reading comprehension immediately after the programme (t2) with

effects maintained at 4-month follow-up (t3). Although the sample was selected as the lowest-performing readers within each class, there was substantial variation in baseline reading ability (including a minority scoring in the average range on

14697610, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpp.70159 by Bangor University Main Library, Wiley Online Library on [14/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

standardised word reading). Homogeneity-of-slopes tests provided no evidence that intervention effects varied as a function of initial ability.

The mediation analyses help to clarify the mechanisms underlying the sustained improvements in reading comprehension at follow-up. In the temporally lagged model, the t3 difference in comprehension was statistically accounted for, in part, by earlier gains in word-level literacy at t2, consistent with the central role of word reading in supporting later comprehension. In an additional exploratory parallel-mediation model, posttest gains in taught vocabulary knowledge also showed an independent indirect pathway to later comprehension alongside word-level literacy. However, the indirect pathway via word-level literacy was stronger than the vocabulary pathway.

It is also important to note that the two mediators differed in their proximity to programme content: the vocabulary factor indexed knowledge of words explicitly taught during RILL (near transfer), whereas the word-level literacy factor was derived from broader decoding/word-reading measures. This difference in measurement scope likely contributed to the relatively larger posttest gain in taught vocabulary, while the larger indirect pathway via word-level literacy reflects the stronger association between general word-level skill and later reading comprehension in the model.

This pattern reflects the stronger association between word-level literacy and later comprehension in the model, and suggests that sustained comprehension benefits may be driven primarily by improvements in word-level skills (Perfetti & Stafura, 2014; Tilstra et al., 2009), with an additional (smaller and more uncertain) contribution from taught vocabulary knowledge. Together, the findings are consistent with multicomponent accounts of reading development in which both word-level and language-related learning support comprehension, but with word-level constraints remaining particularly influential in struggling readers (Gough & Tunmer, 1986; Snowling & Hulme, 2025).

These findings extend the Simple View of Reading (Gough & Tunmer, 1986) by providing causal evidence that decoding continues to constrain comprehension development in struggling readers beyond the early grades. They also align with the Reading is Language model (Snowling & Hulme, 2025), that emphasises cascading and reciprocal relations among reading components: our results provide support for an indirect effect of earlier decoding to later comprehension as postulated by the model.

Practically, the data indicate that sustained gains in comprehension among older, struggling readers can be achieved when interventions produce reliable improvements in word reading. This underlines the importance of explicit decoding instruction beyond the earliest grades and supports integrating decoding and comprehension activities so that children

can repeatedly apply new decoding skills in meaningful contexts. The observed strengthening of comprehension at follow-up—rather than decay—is encouraging for educators and policymakers because durable effects are rare in remediation studies and are critical for lasting academic impact.

### Limitations

One limitation of the study is the absence of repeated, multidimensional measures of oral language (e.g. grammatical or narrative skills), which limits our ability to comprehensively test language-based mechanisms alongside decoding. Although we included a measure of taught vocabulary, this captures only a narrow aspect of oral language and may reflect individual differences in language learning as well as intervention effects. Future studies incorporating richer language measures are needed to clarify these mechanisms.

While randomisation produced comparable groups at baseline and models adjusted for pretest performance, we did not include additional covariates such as socioeconomic status. Although not required for unbiased estimation in an RCT, such variables could improve precision and allow tests of moderation. Finally, the follow-up period was relatively short (4 months), reflecting practical constraints of the school year; longer-term follow-up is needed to determine whether gains persist, widen or attenuate over time.

### Conclusion

This study is among the first RCTs to embed a temporally lagged mediation design to test whether postintervention improvements in word-level literacy contribute to later gains in reading comprehension in older struggling readers. RILL produced immediate and sustained improvements in word-level literacy and reading comprehension, and the follow-up comprehension effect was primarily accounted for by earlier gains in word-level literacy at posttest, with little evidence of a remaining direct effect of the intervention on comprehension. Our findings provide support for a causal effect of word-level literacy skills on later reading comprehension (as postulated by the Simple View of Reading (Gough & Tunmer, 1986) and Reading is Language model (Snowling & Hulme, 2025)). Practically, the results underscore the value of continuing explicit word-reading instruction beyond the earliest grades and integrating it with meaningful text-based and language-rich activities to secure durable improvements in reading comprehension for struggling readers.

### Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

**Appendix S1.** Word-level literacy invariance model fit results.

**Table S1.** Configural, Metric and Scalar Model Fit Estimates Testing Word-level Literacy Invariance between Pretest and Posttest.

**Table S2.** Configural, Metric and Scalar Model Fit Estimates Testing Word-level Literacy Invariance between Pretest and Follow-up.

**Appendix S2.** Reading comprehension invariance model fit results.

**Table S3.** Configural, Metric and Scalar Model Fit Estimates Testing Reading Comprehension Invariance between Pretest and Posttest.

**Table S4.** Configural, Metric and Scalar Model Fit Estimates Testing Reading Comprehension Invariance between Pretest and Follow-up.

**Appendix S3.** Vocabulary invariance model fit results.

**Table S5.** Configural, Metric and Scalar Model Fit Estimates Testing Vocabulary Invariance between Pretest and Posttest.

**Table S6.** Configural, Metric and Scalar Model Fit Estimates Testing Vocabulary Invariance between Pretest and Follow-up.

## Acknowledgements

This project was funded by the Nuffield Foundation (grant reference: EDO/FR-000023452), but the views expressed are those of the authors and not necessarily those of the Foundation. The authors would like to thank James Hodgetts and Carly de Luca for their research assistance. The authors have declared that they have no competing or potential conflicts of interest.

## Data availability statement

The data that support the findings of this study are openly available in OSF at: [https://osf.io/5vmqj/?view\\_only=92f4538009ca4657ab08975d03b364de](https://osf.io/5vmqj/?view_only=92f4538009ca4657ab08975d03b364de).

## Correspondence

Manon Jones, Department of Psychology, Bangor University, Brigantia Building, Penrallt Road, Bangor, Gwynedd, LL57 2AS, UK; Email: [manon.jones@bangor.ac.uk](mailto:manon.jones@bangor.ac.uk)

## Key points

### What's known?

Reading comprehension depends on both accurate word reading and oral language skills. Children who struggle with decoding often continue to experience difficulties with comprehension into the later primary years.

### What's new?

This RCT demonstrates that a language-rich, multicomponent literacy intervention produces immediate improvements in word reading, vocabulary and reading comprehension, with later comprehension gains mediated via improvements in both decoding and vocabulary.

### What's relevant?

Our findings highlight that struggling readers in the later primary years continue to experience a decoding bottleneck but can still make meaningful progress. The intervention tested here is scalable, low-cost and delivered by teaching assistants, underscoring its potential for wider educational practice and policy.

## References

- Adlof, S.M., Catts, H.W., & Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *Journal of Learning Disabilities, 43*, 332–345.
- Ahmed, Y., Miciak, J., Taylor, W.P., & Francis, D.J. (2022). Structure altering effects of a multicomponent reading intervention: An application of the direct and inferential mediation (DIME) model of reading comprehension in upper elementary grades. *Journal of Learning Disabilities, 55*, 58–78.
- Anwyl-Irvine, A.L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J.K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*, 388–407.
- Beck, I.L., McKeown, M.G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary in-struction*. New York: Guilford.
- Beck, I.L., Perfetti, C.A., & McKeown, M.G. (1982). Effects of long-term vocabulary in-struction on lexical access and reading com-prehension. *Journal of Educational Psychology, 74*, 506–521.
- Bowyer-Crane, C., Snowling, M.J., Duff, F.J., Fieldsend, E., Carroll, J.M., Miles, J., ... & Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry, 49*, 422–432.
- Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*, 1398–1407.
- Caravolas, M., Lervåg, A., Mikulajová, M., Defior, S., Seidlová-Málková, G., & Hulme, C. (2019). A cross-linguistic, longitudinal study of the foundations of decoding and reading comprehension ability. *Scientific Studies of Reading, 23*, 386–402.
- Catts, H.W. (2003). Language basis of reading disabilities and implications for early identification and remediation. *Reading Psychology, 24*, 223–246.

- Catts, H.W., Adlof, S.M., & Weismer, S.E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*, 278–293.
- Clarke, P.J., Snowling, M.J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*, 1106–1116.
- de Jong, P.F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading, 6*, 51–77.
- Downing, C., Evans-Jones, G., Calabrich, S.L., Wynne, C., Cartin, R., Dunton, J., ... & Jones, M. (2025). Literacy instruction from afar: Evidence for the effectiveness of a remotely delivered language-rich reading programme. *Reading and Writing, 38*, 77–91.
- Duff, F.J., Fieldsend, E., Bowyer-Crane, C., Hulme, C., Smith, G., Gibbs, S., & Snowling, M.J. (2008). Reading with vocabulary intervention: Evaluation of an instruction for children with poor response to reading intervention. *Journal of Research in Reading, 31*, 319–336.
- Ehri, L.C., Nunes, S.R., Willows, D.M., Schuster, B.V., Yaghouh-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*, 250–287.
- Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review, 23*, 553–576.
- Fricke, S., Bowyer-Crane, C., Haley, A.J., Hulme, C., & Snowling, M.J. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry, 54*, 280–290.
- García, J.R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research, 84*, 74–111.
- Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6–10.
- Hatcher, P.J., Hulme, C., & Ellis, A.W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65*, 41–57.
- Hatcher, P.J., Hulme, C., & Snowling, M.J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry, 45*, 338–358.
- Hulme, C., Bowyer-Crane, C., Carroll, J.M., Duff, F.J., & Snowling, M.J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science, 23*, 572–577.
- Keenan, J.M., Betjemann, R.S., & Olson, R.K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*, 281–300.
- Kendeu, P., Papadopoulou, T., & Kotzopoulou, M. (2013). Evidence for the early emergence of the simple view of reading in a transparent orthography. *Reading and Writing, 26*, 189–204.
- Kim, Y.S., Wagner, R.K., & Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. *Scientific Studies of Reading, 15*, 338–362.
- Kraft, M.A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*, 241–253.
- Latimier, A., Peyre, H., & Ramus, F. (2021). A meta-analytic review of the benefit of spacing out retrieval practice episodes on retention. *Educational Psychology Review, 33*, 959–987.
- Leach, J.M., Scarborough, H.S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology, 95*, 211–224.
- Leppänen, U., Niemi, P., Aunola, K., & Nurmi, J.E. (2004). Development of reading skills among preschool and primary school pupils. *Reading Research Quarterly, 39*, 72–93.
- Lervåg, A. (2021). Is there a core deficit in specific learning disabilities? *Journal of Child Psychology and Psychiatry, 62*, 677–679.
- Lervåg, A., & Hulme, C. (2010). Predicting the growth of early spelling skills: Are there heterogeneous developmental trajectories? *Scientific Studies of Reading, 14*, 485–513.
- Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the developmental relationship between oral language skills and reading comprehension: It's simple, but complex. *Child Development, 89*, 1821–1838.
- Lovett, M.W., Frijters, J.C., Steinbach, K.A., De Palma, M., Lacerenza, L., Wolf, M., ... & Morris, R.D. (2022). Interpreting comprehension outcomes after multiple-component reading intervention for children and adolescents with reading disabilities. *Learning and Individual Differences, 100*, 102224.
- Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin, 140*, 409–433.
- Müller, K., & Brady, S. (2001). Correlates of early reading performance in a transparent orthography. *Reading and Writing, 14*, 757–799.
- Nation, K., Cocksey, J., Taylor, J.S., & Bishop, D.V. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry, 51*, 1031–1039.
- Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing, 23*, 189–208.
- Perfetti, C.A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*, 22–37.
- Roberts, G., Vaughn, S., Fletcher, J., Stuebing, K., & Barth, A. (2013). Effects of a response-based, tiered framework for intervening with struggling readers in middle school. *Reading Research Quarterly, 48*, 237–254.
- Salceda, J.C.R., Alonso, G.A., & Castilla-Earls, A.P. (2014). The simple view of reading in elementary school: A systematic review. *Revista de Logopedia, Foniatria y Audiología, 34*, 17–31.
- Scarborough, H.S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for research in early literacy* (pp. 97–110). New York: Guilford Press.
- Schulz, K.F., Altman, D.G., Moher, D., & Consort Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology, 63*, 834–840.
- Silverman, R.D., Johnson, E., Keane, K., & Khanna, S. (2020). Beyond decoding: A meta-analysis of the effects of language comprehension interventions on K-5 students' language and literacy outcomes. *Reading Research Quarterly, 55*, S207–S233.
- Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica: Rand Corporation.
- Snowling, M.J., & Hulme, C. (2025). The reading is language model: A theoretical framework for language and Reading development and intervention. *Annual Review of Developmental Psychology, 7*, 195–218.
- Snowling, M.J., Stothard, S.E., Clarke, P., Bowyer-Crane, C., Harrington, A., Nation, K., ... & Hulme, C. (2009). *York Assessment of Reading for Comprehension*. GL Assessment.

- Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading, 32*, 383–401.
- Torgesen, J.K., Alexander, A.W., Wagner, R.K., Rashotte, C.A., Voeller, K.K., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34*, 33–58.
- Torppa, M., Tolvanen, A., Poikkeus, A.M., Eklund, K., Lerkkanen, M.K., Leskinen, E., & Lyytinen, H. (2007). Reading development subtypes and their early characteristics. *Annals of Dyslexia, 57*, 3–32.
- Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading, 15*, 8–25.
- Vousden, J.I., Cunningham, A.J., Johnson, H., Waldron, S., Ammi, S., Pillinger, C., ... & Wood, C. (2022). Decoding and comprehension skills mediate the link between a small-group reading programme and English national literacy assessments. *British Journal of Educational Psychology, 92*, 105–130.
- Wilkinson, G.S., & Robertson, G.J. (2017). *WRAT-5: Wide range achievement test professional manual*. United States: Psychological Assessment Resources.
- Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition, 51*, 91–103.

Accepted for publication: 9 April 2026