Spelling Profile Assessment (SPA) a New Spelling Test for English Primary School Children

Presenters:

Dr Georgia Niolaki

g.z.niolaki@bham.ac.uk

The team ☺

Suzanna Presland,
Brenda-Jean Phayre-Mudge,
Heather Marshall,
Claire Hardwire
Cat Philips
Children,
Teachers, SENCos &TAs,
Parents.

Niolaki, G., Terzopoulos, A., Purser, H., Vousden, J., Kohnen, S. and Masterson, J.

Introduction

- **Why spelling?**

- In 2025, only 73% of pupils met the expected standard in GPS; 1% increase from 2024 and 2023 (GOV.UK, 2025).

- Teachers frequently say that they do not know how to support spelling (Esposito et al., 2022).

Introduction

- **Why spelling?**

Weaknesses in spelling negatively impact
- Written expression (Puranik & AlOtaiba, 2012),
- Reading ability (Graham & Santangelo, 2014) and
- Academic self-concept (Schuchardt et al., 2015)

- This is particularly pertinent for those with diagnosed reading and spelling dificulties, such as dyslexia, with studies providing evidence of co-occurrence with mental health conditions such as depression and anxiety (Xiao et al., 2023).

Why create a new spelling test (SPA)

- Children use lexical/whole-word, semantic and sublexical processes to spell (Share, 1995; Caravolas et al., 2001; Niolaki et al., 2020; 2023)

- In the past, more evidence from adults than children

- Dual-route model of spelling (e.g. Houghton & Zorzi, 2003, Barry, 1994)

Rationale for test

- **Why create a new spelling test (SPA)**

- Children, as they grow older, also use knowledge about morphemes (Deacon et al., 2009, 2018; Duncan et al., 2009).

- Morphemes are the smallest units of language that convey meaning

- Free morphemes: stand-alone words (run)

- Bound morphemes: cannot stand alone
  Prefix: *re*-start
  Suffix: boy-*s,* meaning-*less*

Rationale for test

- **Why create a new spelling test (SPA)**

- Most tests on spelling do not distinguish between different processes – good, average, or poor speller only

- So current tests cannot guide intervention or instruction effectively, track student learning over time, or determine teaching effectiveness

UNIQUE CONTRIBUTION

- Our test is unique in assessing spelling performance on three different word types:

- <u>Regular, irregular and pseudo</u> (made-up) words

- And the associated processes: *phonology, orthography and morphology (Triple Word Theory, Garcia, Abbott, & Berninger, 2010)*

PART A

Single-word spelling test (SPA)

regular words (<best>), (phonology+ orthography)

irregular words (<yacht>), (orthography)

pseudo-words (made-up words <ig>) (phonology)
'controlling for prior word specific knowledge'

# UNIQUE CONTRIBUTION

PART B

Morph_Test

morphological rules to spell single words (kicked)

and pseudo-words (vicked)

- Matched on a large number of different psycholinguistic variables

Bespoke spelling test, 'SPA'

- **Pilot phase** (awarded by British Academy 2017-2019 and British Psychological Society research assistantship 2020-2021) (Niolaki et al., 2019 ECDP; Vousden et al., 2019, FRLL)

  - We collected data from:

  - 6 primary schools in the UK N=641, boys=310
  - 36 regular/irregular, 34 pseudowords (Rasch modelling)
  - differentiate between effective and ineffective items
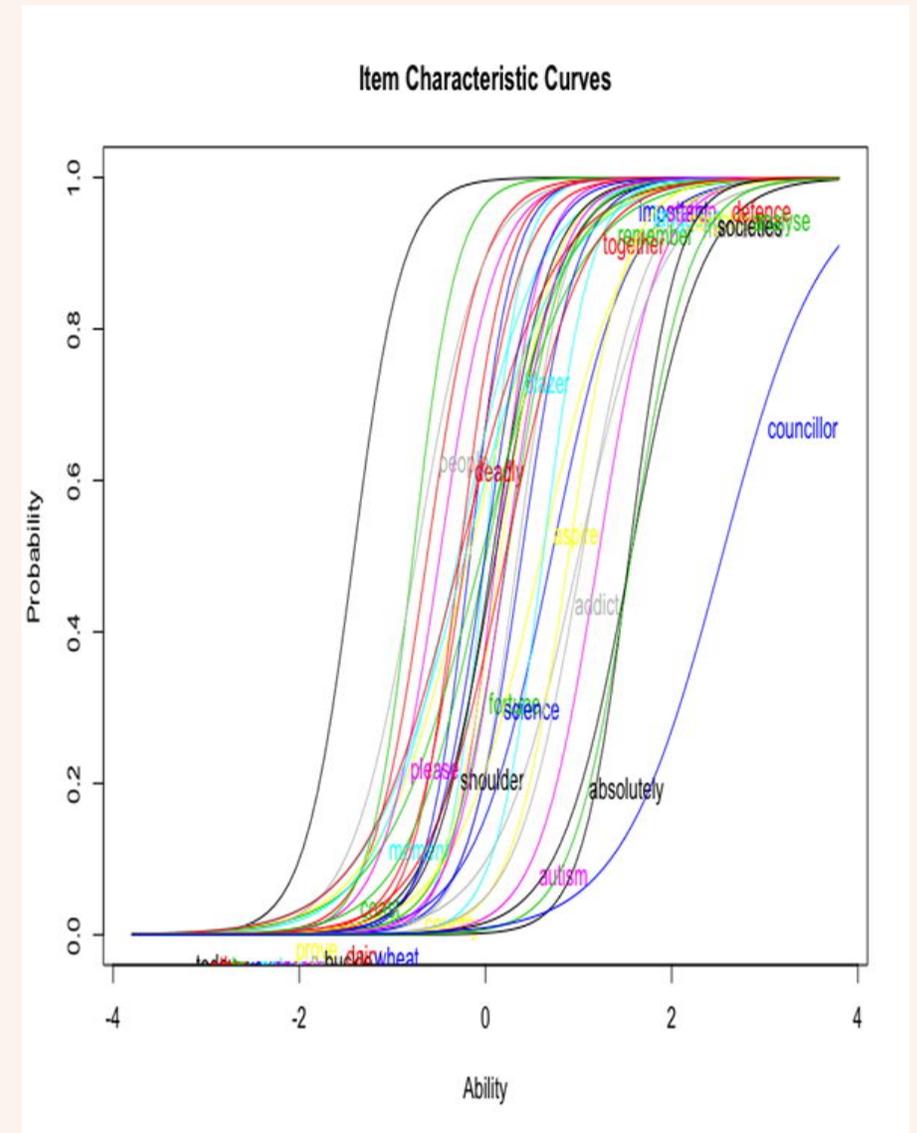
Test Construction SPA

- 29 Regular words

- 29 Irregular Words

- 29 Nonwords/pseudowords

(short, medium, long ~ length & low, medium, high ~ frequency)

factors: length, frequency[a], and phonological neighbours[b] (K-W T, p>.05)

a Van Heuven, W.J.B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67, 1176-1190.*
b Davis, C.J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behaviour Research Methods, 37*, 65-70

Bespoke spelling test, 'SPA'

- *Psycholinguistic characteristics (final phase)*

	Log_ CPWD$_a$	CBeebies$_b$	CBBC$_b$	Zipfreq$_b$	NLetters	NPhonemes	NSyllables	CD zip$_b$	Phon N$_c$
Irregular items	2.02(1.3)	4.3 (1.4)	4.6 (1.3)	4.7 (1.3)	6.5 (1.8)	4.9 (1.6)	2.0 (.86)	.30(.34)	5.5 (5.7)
Regular items	1.5(.77)	4.1 (.94)	4.2 (.96)	4.3 (.95)	6.2 (2.0)	5.43 (1.77)	1.7 (.79)	.18(.22)	6.3 (6.8)
pseudoword	-	-	-	-	6.5 (2.0)	5.6 (1.9)	1.8 (.8)		
Kruskal_Wallis H	.66	.58	3.1	1.8	55	1.5	3.2	2.2	.01
p	.41	.44	.08	.17	.75	.47	.19	.13	.9

Note: a= Masterson et al. (2010); b=Van Heuven et al. (2014); c=Davis (2005)

Pseudowords SPA

- Splitting regular words and recombining

- e.g.

yesterday property -> yesperty

yesterday property -> proterday

- Preserves some psycholinguistic properties for matching to the regular/irregular

Sample

- 65 classes in primary schools in the UK: Birmingham, Midlands, Somerset, London, Bristol, Gloucestershire, Forest of Dean, Bath

N= 1,052	Type	Size
Schools (N=22)	Rural/Urban state, CoE, Independent (6%)	150-500

Yr R	Yr 1	Yr 2	Yr 3	Yr 4	Yr 5	Yr 6
146 f=42% we=75% EAL=24% SEN=8%	60 f=43% we=68% EAL=19% SEN=15%	127 f=50% we=50% EAL=50% SEN=4%	237 f=51% we=75% EAL=22% SEN=6%	155 f=58% we=68% EAL=28% SEN=4%	166 f=48% we=71% EAL=25% SEN=5%	161 f=46% we= 43% EAL=52% SEN=8%

Test Construction Morph-Test

- 20 Real words (20 affixes/ 7 suffixation rules) <u>pollut</u>ion (2 credits)

- 20 Nonwords/pseudowords (20 affixes/ 7 suffixation rules) <u>swit</u>ting (2 credits)

On average medium to long ~ length & on average high ~ frequency)

Factors: length, frequency$_a$, and orthographic neighbours$_b$ (K-W T, p>.05 & Wilcoxon (Z) p>.05)

Note: a= Masterson et al. (2010); b=Davis (2005)

Test Construction Morph-Test

Real words	Log_CPWD$_a$	N Letters	N Phonemes	N Syllables	Orthog_N$_b$
R Yr (2 items)	4.4 (.21)	6 (.0)	5 (1.4)	1.5 (.7)	7.5 (2.1)
Yr 1 (3 items)	4.4(1.8)	5.3 (1.1)	4.6 (.57)	1.6 (.57)	4.6 (8.1)
Yr 2 (3 items)	4.1 (.89)	6.6 (1.1)	6 (1)	2.3 (.5)	.66 (1.1)
Yr 3 (3 items)	2.6 (1.2)	8.6 (.57)	6.6 (1.1)	2.6 (.57	.00 (.00)
Yr 4 (4 items)	3.7 (1.5)	6.5 (1.7)	5.2 (.5)	2 (.81)	1.5 (1.9)
Yr 5 (3 items)	4 (.52	7.6 (3.2)	6.3 (2.1)	2.3 (1.1)	5.6 (8.9)
Yr 6 (2 items)	4.1 (.58)	9 (1.4)	7 (1.4)	3 (1.4)	1 (1.4)
Mean (20 items)	3.9 (1.1)	7 (1.8)	5.8 (1.2)	2.2 (.83)	2.8 (4.8)
Kruskal_Wallis H real words across year_groups	5.1	8.01	7.9	5.5	6.3
p	.52	.23	.24	.48	.39
Pseudowords Mean (20 items)	-	7 (1.8)	6 (1.6)	2.2 (.83)	2.8 (4.5)
Wilcoxon (Z) pseudowords vs. real words	-	1	1.1	.0	.24
p	-	.31	.23	1	.80

Note: a= Masterson et al. (2010); b=Davis (2005)

Pseudowords Morph_Test

- Based on the real words (orthographically appropriate made-up words)

- e.g.

  Kicked-> Vicked

  retell -> remell

- Preserves some psycholinguistic properties for matching to the real word

Morph_Test example items

			Affix spellings	Suffixation rules
1	**Kick****ed**	He **kicked** the ball yesterday and scored two goals. **Kicked**	1 for -ed correct	--
2	**fast****er**	I can run fast, but he can run **faster** than me. **Faster**	1 for -er correct	--
1	**Vick****ed**	They were able to vick one goal today; Yesterday they **vicked** two goals. **Vicked**	1 for -ed correct	**--**
2	**Dast****er**	This toy is very dast, but my friend has a **daster** toy. **Daster**.	1 for -ed correct	--

Sample

- 36 classes in primary schools in the UK: Somerset, Bristol, Gloucestershire, Bath …

N=409	Type	Size
Schools (N=5)	Rural/Urban, state, CoF, Independent (8.6%)	150-500

Yr R	Yr 1	Yr 2	Yr 3	Yr 4	Yr 5	Yr 6
45 f=33% we=97% EAL=0% SEN=11%	39 f=43% we=100% EAL=0% SEN=23%	61 f=44% we=97% EAL=3% SEN=8%	74 f=41% we=91% EAL=8% SEN=17%	69 f=65% we=88% EAL=14% SEN=8%	70 f=42% we=90% EAL=7% SEN=10%	51 f=51% we=90% EAL=0% SEN=9%

Measures-Procedure

- *Literacy-related variables*
  - SPA
  - Morph_Test
  - Spelling (WIAT 3) ~α= The average reliability reported for 5 to 12-year-olds is .94 *s.d.*=.02

- Ethical approval was obtained from the University's Ethics Committee before data collection began. Following the school heads' agreement to testing, parents/carers were asked to consent, and participating children were asked for assent.

- Data were collected between May and June by trained and paid research assistants (teachers in primary schools).

Mean accuracy of SPA, Morph test and WIAT3 spelling test per year group (standard deviations are in parenthesis)

|	SPA			Morph test		WIAT 3
Year_group	Irregular words (n 29)	Regular words (n 29)	Pseudowords (n 29)	Real words (n 27)	Pseudowords (n 27)	(n 63)
R Yr	.54 (1.4)	2.16 (2.8	2.23 (3.4)	1.07 (1.4)	.48 (1.1)	5.56 (2.8)
Yr 1	4.51 (3.6)	6.19 (6.7)	6.95 (7.8)	4.21 (3.3)	3.23 (2.5)	11.73 (4.7)
Yr 2	6.52 (5)	12.53 (6.8)	13.22 (8.3)	6.91(5.6)	5.24 (4.6)	16.22 (5.8)
Yr 3	12.95 (6.5)	18.88 (6.6)	16.19 (6.3)	13.33 (5.9)	10 (5)	21.87 (7.5)
Yr 4	17.8 (5.6)	23.04(5.1)	19.46 (6.4)	15.32 (5.1)	11.65 (5.3)	25.37(9.4)
Yr 5	19.82 (5.2)	24.39 (5)	21.87 (5.1)	16.41(5.4)	14.1 (5.1)	31.38(9.1)
Yr 6	21.98 (5.1)	25.52 (3.9)	21.96 (4.5)	20.22 (4.1)	17.88 (4.2)	36.71 (10.4)
Total	**12.78 (8.9)**	**17.21 (9.7)**	**15.36 (9)**	**11.77 (7.7)**	**9.46 (6.9)**	**21.58 (11.9)**

Results

SPA Mean accuracy spelling for regular, irregular and pseudowords and interaction with Year_group.
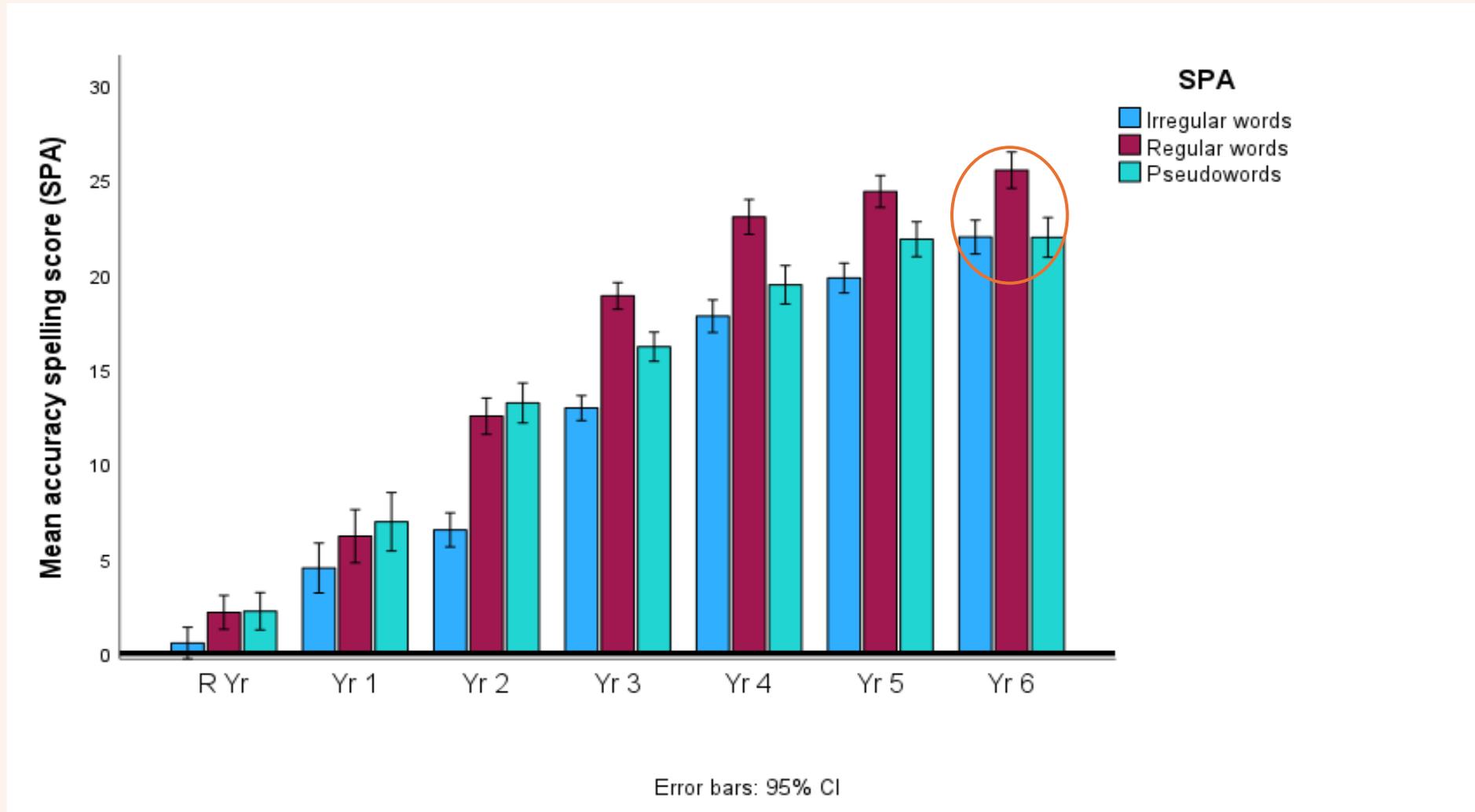
Results

Morph test mean accuracy spelling for real words and pseudowords and interaction with Year_group.
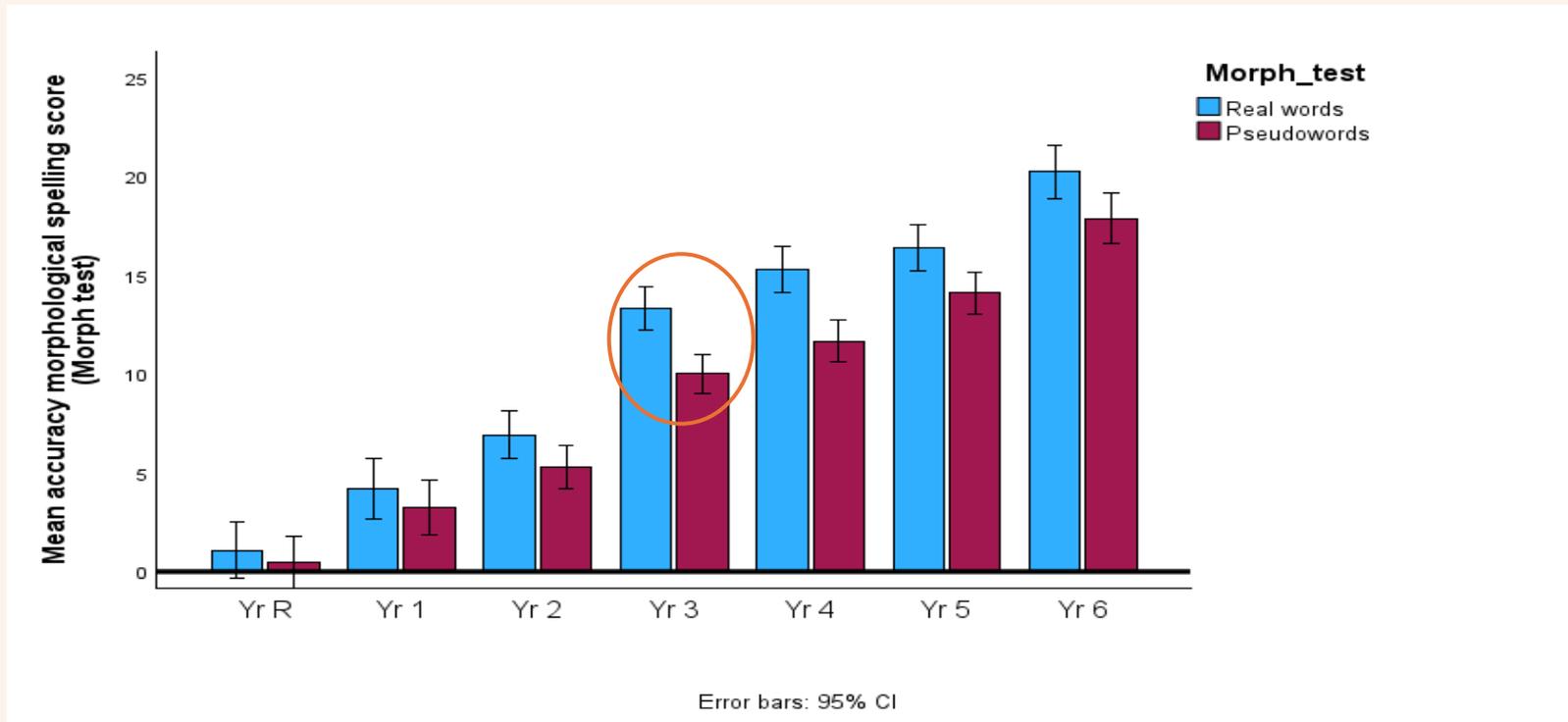
Reliability

- Cronbach's alpha SPA

	Yr R			Yr 1			Yr 2			Yr 3			Yr 4			Yr 5			Yr 6		
	Ir	Rg	Ps	Ir	Rg	Ps	Ir	Rg	Ps	Ir	Rg	Ps	Ir	Rg	Ps	Ir	Rg	Ps	Ir	Rg	Ps
α	.93	.98	.98	.95	.93	.92	.95	.92	.88	.98	.97	.96	.99	.98	.92	.99	.73	.94	.97	.92	.90

- Total sample α = irregular words .99, regular words .97, pseudowords .97

Reliability

- Cronbach's alpha Morph_Test

	Yr R		Yr 1		Yr 2		Yr 3		Yr 4		Yr 5		Yr 6	
	Rw	Ps	Rw	Ps	Rw	Ps	Rw	Ps	Rw	Ps	Rw	Ps	Rw	Ps
α	.85	.46	.92	.93	.91	.91	.94	.93	.94	.90	.90	.85	.90	.84

- Total sample $\quad\quad\quad \alpha$ = real words .98, pseudowords .97

*Validity Pearson correlation denoted as **r***

- Excellent criterion and convergent validity

	1. Spelling (WIAT 3)	2. SPA (Irregular)	3. SPA (Regular)	4. SPA (Pseudoword)	5. SPA total	6. Morp_test real word	7. Morp_test pseudo word	8. Morph_Test total
2	0.58****	-						
3	0.52****	0.78****	-					
4	0.52****	0.53****	0.75****	-				
5	0.61****	0.85****	0.94****	0.87****	-			
6	0.50****	0.67****	0.68****	0.61****	0.73****	-		
7	0.54****	0.46****	0.56****	0.61****	0.61****	0.74****	-	
8	0.56****	0.59****	0.65****	0.66****	0.72****	0.92****	0.95****	-

raw scores

Example differentiation

Participant	SPA test SS				Morph_Test			WIAT-3
	Irregular	Regular	Pseudowords	SPA CS	Realwrs	Pseudow	Morph_test CS	
Male A Y3 (8.4 years)	82	83	85	83	80	90	85	77
Female A Y3 (8.3 years)	74	77	96	81	87	93	90	80
Female B Y3 (8.5 years)	82	83	107	91	89	94	95	94
Male B Y3 (8.3 years)	84	104	107	98	103	110	108	98

Male Y3 A (8:04 years) ADHD

Irregular words	
other	over
please	plese
deadly	deley
dairy	derey
young	yun
nature	nacher
remember	✔
shoulder	sholder
search	serch

Morph_Test real words

Kicked	kite
faster	✔
man's	✔
retell	✔
famous	famos
families	fameley
mislay	✔
oldest	oldis
amazement	amasmont
harmless	haremlise
pollution	poloshon
danced	douns
wonderful	wondfall

Male Y3 B (8:03 years) FSM irregular word spelling

was	✔	shoulder	sholder
said	✔	search	serch
have	✔	country	contry
come	✔	station	staiton
people	✔	fortune	forchun
other	✔	flavour	flaver
please	Pleas	science	since
deadly	✔	aspire	aspiyer
dairy	dariy	language	langwich
young	yung	wheat	weat
nature	nacher	neighbour	nabor
remember	reamember	marvellous	marvilus

Conclusion

- A better understanding of spelling development can be achieved using tools that distinguish between regular and irregular words and pseudowords, as well as capture morphological rules.

- Analysis indicates that the test is a valid and reliable measurement of spelling skill capturing phonological, orthographic and morphological rules.

- Targeted intervention can be put in place when we know the learner's spelling needs.

Thank you

- ☺

- Time for questions!