

Technical information for NFER tests in spelling Suite 2

Centre for Assessment

Published in January 2017

By the National Foundation for Educational Research,
The Mere, Upton Park, Slough, Berkshire SL1 2DQ

www.nfer.ac.uk

© 2017 National Foundation for Educational Research
Registered Charity No. 313392

ISBN 978-1-911039-23-5

How to cite this publication:

National Foundation for Educational Research (2017). *Technical Information for NFER Tests in Spelling Suite 2*. Slough: NFER.

Contents

1	Introduction	1
2	Early development of items	2
3	Sample characteristics	3
4	Whole test functioning	6
5	Item level functioning	8
	Item level statistics	8
	Differential item functioning	8
6	Test outcomes	10

1 Introduction

In January 2015, NFER released a set of spelling assessments and an accompanying teacher guide for year 5. This was followed by a set for each of year 3 and year 4 in January 2016. The finalised materials consist of three standardised spelling tests of equivalent difficulty for each year group. The tests are aligned to the 2014 National Curriculum and are meant to be administered by teachers at three points in the academic year for the purpose of measuring pupil progress. The teacher guide provides information about commonly occurring errors and patterns that occur in misspellings.

2 Development of items

The development of NFER spelling tests for each year group involved both qualitative and quantitative work with schools. Early activities involved creating word lists, mapped to the National Curriculum, development of items and expert reviews.

Each spelling test was comprised of twenty four stand-alone sentences (items) with each sentence testing the spelling of one target word. Target words were selected based on the new (2014) National Curriculum Programme of Study for spelling. For the year 5 test, the majority of target words were selected based on the statutory requirements in the year 5 and 6 Programme of Study (PoS), with some coverage of the PoS for spelling in years below year 5 and some extension into spelling coverage beyond year 6. Similarly, for the year 3 and year 4 tests, the majority of target words were selected based on the statutory requirements in the year 3 and 4 PoS, with some coverage of the PoS for spelling in years below year 3 and some extension into spelling coverage beyond year 4.

As part of the large scale standardisations, eight spelling tests for each year group were trialled in schools. Each target word was trialled in two different versions of the tests which helped to gather more reliable data on each item's performance.

During the trial, teachers' views were gathered through a short questionnaire. This enabled us to estimate the demand of materials, establish pupils' level of engagement with them, and also to collect feedback from teachers.

The test papers were returned from the schools and marked by NFER researchers. The summary for whole test functioning statistics is shown in section 4.

The target words were marked using coding that enabled analysis of common and recurring spelling errors. To determine which target words to include in the final versions of the tests, the item functioning of each target word was analysed statistically and low functioning target words were discarded. Three final tests per year group were then constructed based on the statistical performance and suitable coverage of the curriculum.

3 Sample characteristics

A large scale standardisation trial was conducted in June 2014 for year 5 and in June 2015 for years 3 and 4. The standardisation sample was stratified according to the following characteristics:

- KS2 overall performance band 2013 (average point score)
- Primary school type
- Region: government office region

The characteristics of the sample of schools involved in the standardisation of the NFER spelling tests can be compared with the characteristics of all maintained schools in England with pupils of the relevant age on roll. The samples are compared in Tables 1 – 3 below. Significance tests were conducted on the samples which show that the achieved samples are representative of the national picture. The gender breakdown of the sample is shown in Table 4.

Table 1: Representation of the sample at school level - Year 3 spelling

		population		sample	
		Number	%	Number	%
Primary school type	Primary/Combined	11,057	70	24	59
	Junior	939	6	4	10
	Middle	16	<1	0	0
	Independent schools	1,242	8	3	7
	Other type	2,522	16	10	24
Region	North	4,906	31	11	27
	Midlands	4,810	30	14	34
	South	6,060	38	16	39
KS2 overall performance band 2013 (av. point score)	Lowest 20%	2,560	16	5	12
	2nd lowest 20%	2,389	15	6	15
	Middle 20%	2,616	17	8	20
	2nd highest 20%	2,440	15	11	27
	Highest 20%	2,954	19	7	17
	Missing	2,817	18	4	10
Total schools		15,776	100	41	100

Since percentages are rounded to the nearest integer, they may not always sum to 100.

Table 2: Representation of the sample at school level - Year 4 spelling

		population		sample	
		Number	%	Number	%
Primary school type	Primary/Combined	10,916	70	29	62
	Junior	941	6	2	4
	Middle	27	<1	0	0
	Independent schools	1,255	8	4	9
	Other type	2,479	16	12	26
Region	North	4,895	31	14	30
	Midlands	4,711	30	15	32
	South	6,012	38	18	38
KS2 overall performance band 2013 (av. point score)	Lowest 20%	2,531	16	4	9
	2nd lowest 20%	2,369	15	8	17
	Middle 20%	2,605	17	6	13
	2nd highest 20%	2,420	15	10	21
	Highest 20%	2,946	19	15	32
	Missing	2,747	18	4	9
Total schools		15,618	100	47	100

Since percentages are rounded to the nearest integer, they may not always sum to 100.

Table 3: Representation of the sample at school level - Year 5 spelling

		population		sample	
		Number	%	Number	%
Primary school type	Infant/First	12	<1	0	0
	Primary/Combined	11,247	77	37	90
	Junior	987	7	0	0
	Middle	183	1	1	2
	Other type	2,189	15	1	2
	Missing	0	0	2	5
Region	North	4,750	32	15	37
	Midlands	4,522	31	6	15
	South	5,346	37	18	44
	Missing	0	0	2	5
KS2 overall performance band 2013 (av. point score)	Lowest 20%	2,622	18	10	24
	2nd lowest 20%	2,430	17	7	17
	Middle 20%	2,661	18	8	20
	2nd highest 20%	2,442	17	5	12
	Highest 20%	2,954	20	8	20
	Missing	1,509	10	3	7
Total schools		14,618	100	41	100

Since percentages are rounded to the nearest integer, they may not always sum to 100.

Table 4: Representation of the sample at pupil level: gender

	population	Y3 sample		Y4 sample		Y5 sample	
	%	Number	%	Number	%	Number	%
Boys	51	742	51	766	51	688	49
Girls	49	711	49	748	49	688	49
Missing	0	0	0	0	0	17	1
Total	100	1453	100	1514	100	1393	100

Since percentages are rounded to the nearest integer, they may not always sum to 100.

4 Whole test functioning

The following tables provide information on the overall functioning of each test trialled, separately by year group. Eight test versions were trialled per year group.

Table 5: Whole test functioning for year 3, trial versions

	Trial version							
	1	2	3	4	5	6	7	8
Sample <i>n</i>	181	226	188	184	159	130	168	217
Reliability (Cronbach's alpha)	0.936	0.927	0.912	0.923	0.910	0.935	0.916	0.919
Maximum score	24	24	24	24	24	24	24	24
Mean	11.52	11.25	12.18	12.28	10.76	10.52	9.46	10.79
Standard deviation	7.16	6.50	6.10	6.71	6.23	6.92	6.18	6.55

Table 6: Whole test functioning for year 4, trial versions

	Trial version							
	1	2	3	4	5	6	7	8
Sample <i>n</i>	204	236	174	183	209	210	216	82
Reliability (Cronbach's alpha)	0.902	0.907	0.894	0.902	0.890	0.898	0.918	0.902
Maximum score	24	24	24	24	24	24	24	24
Mean	10.91	8.58	7.46	8.49	8.72	8.04	8.68	8.40
Standard deviation	6.23	6.03	5.38	5.85	5.48	5.64	6.20	5.71

Table 7: Whole test functioning for year 5, trial versions

	Trial version							
	1	2	3	4	5	6	7	8
Sample <i>n</i>	135	178	186	176	169	183	182	184
Reliability (Cronbach's alpha)	0.920	0.907	0.916	0.900	0.896	0.884	0.907	0.883
Maximum score	24	24	24	24	24	24	24	24
Mean	8.81	9.66	14.36	10.45	10.71	10.91	8.16	8.83
Standard deviation	6.37	6.01	6.16	5.73	5.70	5.36	5.71	5.33

The following tables provide information on the overall functioning of the final tests.

Table 8: Whole test functioning for year 3, final tests

	Test version		
	1	2	3
Sample n*	1453	1453	1453
Reliability (KR21)	0.923	0.920	0.913
Maximum score	24	24	24
Mean	11.86	12.07	12.05
Standard deviation	7.22	7.11	6.93

Table 9: Whole test functioning for year 4, final tests

	Test version		
	1	2	3
Sample n*	1514	1514	1514
Reliability (KR21)	0.907	0.913	0.895
Maximum score	24	24	24
Mean	10.02	10.82	10.71
Standard deviation	6.68	6.89	6.46

Table 10: Whole test functioning for year 5, final tests

	Test version		
	1	2	3
Sample n*	1393	1393	1393
Reliability (KR21)	0.881	0.906	0.880
Maximum score	24	24	24
Mean	11.10	11.60	10.83
Standard deviation	6.20	6.76	6.16

**the total number of pupils in the trial sample for this year group*

5 Item level functioning

Item level statistics

Information about item functioning is available in the 'question grid' spreadsheets. These are downloadable from the NFER portal for purchasers of the tests.

Differential item functioning

Differential item functioning (DIF) analysis for gender was carried out to identify observed differences in performance on each test. Differential item functioning identifies particular items for which two groups (in this case girls and boys) perform differently above and beyond the disparity in their achievement on the test as a whole.

The following tables present the outcomes of the DIF analyses, showing the questions where statistical differences between groups have been identified. It is important to note that these statistical differences do not necessarily indicate that a particular question is globally biased towards one group or the other, but may reflect genuine differences in performance for this sample of pupils.

All items (target words) were trialled in two different test versions. However, in most cases where a DIF was identified for target words, the DIF only applied to one version of the test where the item appeared, but not the other test version.

Table 11: Differential item functioning for gender (Years 3, 4 and 5)

	Item (target word)	Test	DIF favours	
Year 3	address	C	Boys	**
	bought	C	Boys	*
	delighted	B	Girls	*
	editing	C	Girls	*
	slice	C	Boys	**
	special	B	Boys	**
	stories	C	Boys	**
	wonder	A	Boys	*
Year 4	caution	C	Girls	*
	chalk	C	Boys	**
	collector	C	Girls	*
	dangerously	A	Boys	*
	entertaining	B	Boys	*
	equipment	C	Girls	**
	heavier	C	Boys	**
	inspiration	A	Girls	*
	lightning	B	Boys	*
	lorries	A	Boys	*
	prey	A	Girls	*
	remaining	A	Girls	*
	signature	C	Girls	*
	substitute	A	Girls	***
	weight	B	Boys	**
wriggle	A	Boys	*	
Year 5	bruised	B	Girls	**
	comb	B	Girls	**
	fragrance	A	Girls	*
	importance	B	Girls	*
	important	C	Girls	*
	interval	C	Boys	**
	knot	C	Girls	*
	logically	A	Girls	*
	magnificent	B	Boys	*
	misplaced	B	Boys	**
	planning	C	Girls	**
	radiant	B	Boys	**
	reliant	A	Boys	*
	reliant	A	Boys	*

* $p < .05$

** $p < .01$

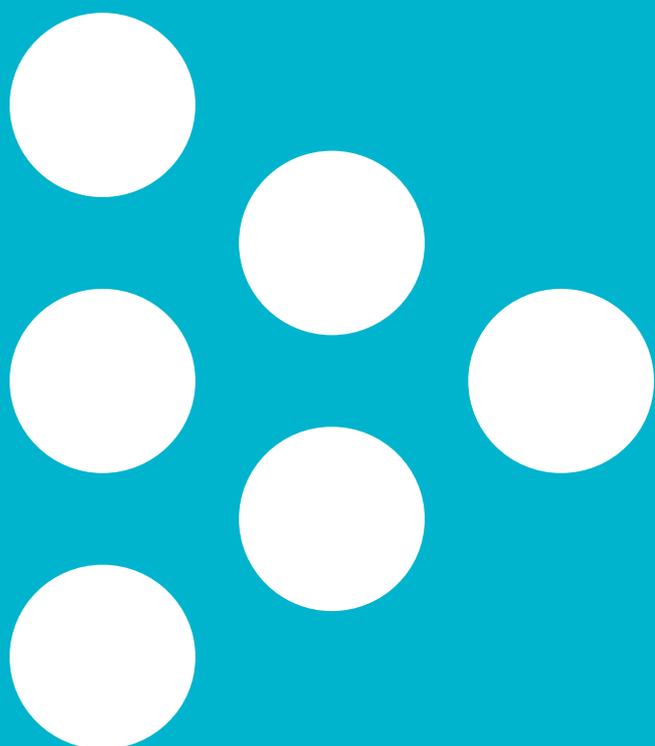
*** $p < .001$

6 Test outcomes

The following outcomes are available from this suite of tests:

- Raw score – the total number of marks attained by each pupil
- Standardised score
- Age-standardised score

More details of each are available in the relevant teacher guide.



Evidence for excellence in education

© National Foundation for Educational Research 2018

All rights reserved. No part of this document may be reproduced or transmitted in any form or by any means, electronic, mechanical, photocopying, or otherwise, without prior written permission of NFER.

The Mere, Upton Park, Slough, Berks SL1 2DQ
T: +44 (0)1753 574123 • F: +44 (0)1753 691632
enquiries@nfer.ac.uk

www.nfer.ac.uk

Product information and prices are correct at the time of printing and are subject to change. Visit our website for latest information.