

Nelson English Language Tests

Nelson English Language Tests Teachers' Book

W S Fowler & Norman Coe

Nelson

Thomas Nelson and Sons Ltd
Lincoln Way Windmill Road
Sunbury-on-Thames Middlesex

P.O. Box 73146 Nairobi Kenya

P.O. Box 943 Kingston Jamaica

Thomas Nelson (Australia) Ltd
19-39 Jeffcott Street West Melbourne
Victoria 3003

Thomas Nelson and Sons (Canada) Ltd
81 Curlew Drive Don Mills Ontario

Thomas Nelson (Nigeria) Ltd
8 Ilupeju Bypass PMB 1303 Ikeja Lagos

© W. S. Fowler and Norman Coe 1976
First published 1976, Reprinted 1977, 1979
0 17 555199 5

All Rights Reserved. No part of this
publication may be reproduced, stored in a
retrieval system, or transmitted, in any
form or by any means, electronic,
mechanical, photocopying, recording or
otherwise, without the prior permission of
Thomas Nelson and Sons Limited.

Phototypeset by Tradespools Ltd, Frome,
Somerset
Printed in Great Britain by
Butler & Tanner Ltd,
Frome and London

Contents

The Test Battery	7
Composition of the tests	7
Pre-testing in field trials	8
Uses of the Test Battery	9
How the tests relate to schools	10
Measurement of progress	11
Administering the tests	12
Relating tests to Cambridge exams	12
Relationship between adjacent levels	13
Types of question at the various levels	15
Keys to all tests	16
Model Answer Sheet	56

Acknowledgements

The authors wish to thank Mr Harold Norminton, Director of the British Institute, and his office and teaching staff, especially Mr John Pidcock and Mr William Wilkinson for their useful comments and suggestions; Miss Ana Lankester for her patience in coping with many drafts; Mr Peter Clements at the Institucion del CIC, and Messrs Seamus Farrell and Ernest Hughes at the Dublin School of English for their co-operation in the early field trials. For their co-operation over the later field trials in Great Britain and Ireland the authors wish to thank principals and teachers too numerous to mention by name. Their kind help is much appreciated.

The Test Battery

The 40 tests in this series make up a battery; there are 4 tests, equivalent in difficulty, at each of ten levels from beginners to advanced. The levels are numbered 050, 100, 150, etc. up to 500. The contents of the tests at any level are related to what an average student can be expected to cover during the corresponding number of hours of study. For example, a 250 test is designed for a student who has studied 250 hours of English.

Students nowadays learn English using methods and materials which presuppose a systematic progression from simpler to more complex grammatical structures. Although there is some slight variation in this respect among teachers and materials, there is in general broad agreement as to what should be taught earlier and what later. The contents of these tests are based on a careful structural progression. Moreover, the tests also take into account the appropriate use of the various structures, which is what underlies all situational and notional syllabuses. This is achieved by carefully contextualising the structures, and in particular by making extensive use of continuous prose. In this way we aim to test both grammatical control and communicative competence. Naturally the latter becomes easier to test at the higher levels; at the lower levels one is often forced to resort to single sentence items.

Composition of the tests

Each of the 40 tests contains 50 items and in every case the students have to choose the correct answer from four choices. At each of the ten levels, therefore, there are 200 different items. The extent to which the items can be different from each other depends, of course, on the quantity of material available to the test constructor. In the early stages, particularly in the case of the 050 and 100 levels, the items on the four tests necessarily show great similarity since there is so little which is available for testing at this stage. From then on the four tests at each level show increasing divergence one from another until – at the 450 and 500 levels – the 200 items are virtually all testing different things.

Whereas there is a great measure of agreement among course writers about the selection and ordering of structures, there is much less agreement on the selection and ordering of vocabulary. Since these tests are not based on any one course book it would clearly be unfair to students – and would, of course, make the tests unreliable – if emphasis were given to the knowledge of separate words. For this reason, the lexical load is light in the earlier levels. However, from the 300 level upwards there are items which test the difference between, for example,

peculiar particular private personal.

The continuous prose passages, which are a feature of all tests from level 150 upwards, are not cloze passages, since the latter are produced by omitting every fifth or sixth word regularly throughout the piece. Instead, we have chosen to test specific words and phrases. We have found these passages invaluable for testing the verb phrase (tense, aspect, modal) since this is the only way to provide sufficient context for the answer to be clear-cut.

Many of the items test two structures at the same time, and this makes the tests as a whole more powerful than they would otherwise be. Here is an example:

Is there the bottle?

A much water into

C many water into

B much water in

D many water in

The student has to solve both the *much/many* and the *in/into* distinctions in order to get the right answer.

In addition there are items in the 400 level and upwards which test control of common idioms. Here is an example:

"He calls in from time to time" means

A He often comes to see us

C He visits us every hour

B He visits us occasionally

D He sometimes comes to stay for a while.

Although these tests are concerned mainly with written skills, there is, from the 250 level upwards, an element of pronunciation control. This takes the form in the levels 250 to 400 of selection of a word which rhymes with a given word. In the 450 and 500 levels the pronunciation items are more complicated and therefore more far-reaching. In each case, there exists a word which has the same pronunciation as the given word but a different spelling. The four choices are definitions of the homonym and of three words pronounced similarly but not identically. Here is an example:

Caught

A something worn

C vehicle pulled by a horse

B child's bed

D place where law cases are heard

Here the student has to work out *coat*, *cot*, *cart* and *court*, and then to know that *court* is pronounced the same as *caught*.

While the tests cover a wide range of items appropriate to the various levels, they clearly cannot include within their scope the measurement of specific skills such as discussion or composition writing. What is measured is the general level of English which forms the basis for these and all other specific uses of the language.

Pre-testing in field trials

In good tests each item must distinguish as powerfully as possible between good and weak students and, ideally, each item must test something different from all the other items on the same test. While test constructors bear both these factors in mind in the original selection of contents, clearly the only reason for maintaining any item in the final version of a test is that field trials have proved that the item does in fact do its job.

All the items in these tests have been carefully pre-tested. The field trials were in two stages: first with students whose native language was usually Spanish, and secondly with students from a wide range of language backgrounds. The data obtained from these trials allowed the constructors to revise or eliminate items which were doubtful or ineffective, and also enabled them to adjust the difficulty of all the tests so as to make the four tests at each level equivalent in difficulty and produce even spacing between the different levels.

Uses of the Test Battery

1 Placement

Students who arrive at a school and need to be placed in an appropriate class should be given a brief interview designed to find out about their previous experience and knowledge of English. The interviewer should then administer a test which he thinks will be at the level of the student. The result of this will quickly confirm or modify the interviewer's assessment. Since all the levels are statistically related to the ones on either side, the interviewer is allowed a margin of error without any serious consequences, because he can immediately relate the score obtained by the student to what the student would have scored if he had taken the test above or below it.

2 Diagnosis

Each of the 40 tests covers a wide range of language items appropriate to the level in question. (For details of the types of question, see Table 2, page 15). Compared with a test which has ten or twelve questions on the same three or four structures, the tests in this battery cover a large proportion of the structures at any level. Thus when the test is administered to a group of students, the teacher obtains a clear overall picture of the weaknesses of his class and consequently has an extremely sound base for the remedial work which his class needs.

While the diagnostic power of the Test Battery will be appreciated by teachers at all levels, it will be found particularly useful in those classes which are preparing for Cambridge examinations. For example, students who are in the final stages of preparation for the First Certificate examination should be set more than one of the tests at the 400 level. The result will show the teacher what remedial work still remains to be covered. The 350 and 500 levels serve a similar function for those students who still have some 50 hours of classes before the First Certificate and Proficiency examinations respectively.

3 Students' progress

Each of the ten levels relates to the ones above and below it, (see Table 1, page 14), thus facilitating the measurement of progress. If a student scores, say, 32 on a 250 level test and then makes normal progress over 50 hours of study, he should then score 37 on a 250 test or 32 on a 300 test. In other words, the progress to be expected from 50 hours of study will normally raise a student's score by five points on an equivalent test or give him the same score on a test from the next level above.

End of course exams should normally consist of more than one type of test and these will usually include a combination of objective and subjective elements. An appropriate test from the Battery can be combined with a second test whose form and content will vary according to the level and aims of the course: listening or reading comprehension, dictation, guided or free composition, and so on. While the subjective test need not be scored on the same scale (0 to 50) as the battery test, nevertheless the range of scores for a group of students should be similar on both tests. This is because the range of scores obtained from any test accurately reflects the range of ability of the students. Consequently, the range on other tests should be similar, while not necessarily giving the same order of merit for each individual student.

For example, if Class A scores on a battery test between 24 and 37 with a concentration around 32, then their other tests out of, say, 25 should range between 12 and 19 with a concentration around 16.

How the tests relate to schools

Clearly the tests will relate differently to different types of school. However, with tests at each of 10 evenly spaced levels from beginners to advanced (almost Cambridge Proficiency standard) every type of school can readily be related to the range of tests. Here are three examples.

Example 1. A school with six levels, taking First Certificate at the end of the fourth and Proficiency at the end of the sixth level.

Class Level	Placement/ Entrance test	Passmark	Progress/ Exit test	Passmark
1	—	—	100	30
2	100	30	200	30
3	200	30	300	30
4	300	30	400	30
			or First Certificate	
5	400	30	500	30
	or First Certificate			
6	500	30	Proficiency	

Example 2. A school with three levels, taking First Certificate at the end of the third level.

Class Level	Placement/ Entrance test	Passmark	Progress/ Exit test	Passmark
1	—	—	100	33*
			or 150	28
2	100	33	250	32
	or 150	28	or 300	27
3	250	32	400	30
	or 300	27	or First Certificate	

*Note: in this example the slight discrepancy in the suggested passmarks is a result of matching 3 class levels with 8 test levels. 3 does not divide into 8 without remainder.

Example 3. A school with 10 class levels which does not accept absolute beginners, and which takes First Certificate after the seventh and Proficiency after the tenth level.

Class Level	Placement/ Entrance test	Passmark	Progress/ Exit test	Passmark
1	100	28	100	32
2	100	32	150	30
3	150	30	200	30
4	200	30	250	30
5	250	30	300	30
6	300	30	350	30
7	350	30	400	30
			or First Certificate	
8	400 or First Certificate	30	450	30
9	450	30	500	30
10	500	30	Proficiency	

In these three examples, as also in the relation of Test Battery levels to any other school system, suitable passmarks suggested here must necessarily be somewhat tentative. Only experience in the actual situation will provide a definite answer for each individual school.

Measurement of progress

With the Test Battery your school has 4 equivalent tests at any level. It is thus a simple matter to compare two parallel classes covering the same material, or to compare any present class with any or all of its predecessors. One therefore has an instrument for measuring the state, or progress of any generation of students and for comparing this with any other so that the school can immediately see whether standards are being maintained or whether they are fluctuating.

Administering the tests

Allow 40–50 minutes to complete each test. Tests should also be quick and problem-free to administer and to score.

One of the main problems facing schools is the question of security – how does one ensure that results are not invalidated because of advance knowledge of the contents or because of cheating? In the first place, the Test Battery is marketed in a form which is unattractive to students since the books contain far more material than will be useful to the student at any one time. Secondly, there are four tests at each level. No student need ever know which of the four he is going to get even if he knows the level in question beforehand. Thirdly, suppose a student who knew the level managed to get hold of all four tests at that level, persuaded someone to tell him the answers to all 200 items, and then completed the test satisfactorily. We would argue that a student who could achieve this had learned enough English – either by going through the tests or otherwise – to pass the test anyway, and thus his score would not be invalidated. Finally, even if students at the end of, say, the third class are normally set a 300 test, the school can surprise them by setting a 250 or a 350 test, and later adjust the scores because of the statistical relationship between adjacent levels.

For placement of individual students who arrive at the school, one should normally hand out in sequence an A paper, then a B, then a C, then a D at each level. Friends applying together thus receive different tests and, even if there is no or minimal invigilation, cannot cheat.

For progress testing of classes within the school, there are the two dangers of cheating during the exam and of telling a subsequent class the answers. Both of these risks are minimised by an ordered distribution of the four papers at each level.

Another major headache connected with many exams derives from the amount of time, and therefore also concentration, that teachers have to give in scoring papers. With the battery tests it is a simple matter to duplicate or photocopy an answer sheet with the same format as the keys given at the back of this Handbook. Quick scoring can then be achieved by comparison with the key. Alternatively, transparent masks to fit over the papers can easily be prepared beforehand. Acetate squares for the overhead projector, for example, are ideal for this purpose. Scoring time can thus be reduced to one or two minutes per paper.

Relating tests to Cambridge exams

The 400 level is equivalent in difficulty to the Cambridge First Certificate. In other words, a student who has not taken the First Certificate may be placed with students who have done so, provided he passes a 400 test.

The 350 level is an accurate predictor for potential candidates for the First Certificate exam. Assuming that the students will have some 50 hours of tuition between the time of enrolment and the date of the exam, then a score of 30 indicates an even chance in the coming exam; a lower score indicates a probable failure and a higher score a probable pass.

The 500 level tests are related to the Proficiency exam as the 350 level is related to the First Certificate, again with 30 as the borderline.

Although the Cambridge exams make considerable use of multiple choice tests, it is of course true that Cambridge candidates also need particular skills, such as the ability to write a composition. These skills are not measured by the Test Battery, but the 350, 400 and 500 levels correlate very highly with the level of language ability required at the stages outlined above.

Relationship between adjacent levels

At each level the four papers A, B, C, and D are equivalent in difficulty. They are designed for a 30 (60%) passmark in all cases, although schools may adjust this in the light of their particular needs. Each level is related to the levels on either side of it as shown in Table 1 (page 14). For example, if a student scores 35 on a 250 test, this is equivalent to his scoring 40 on a 200 test or 30 on a 300 test. Similarly, a score of 27 on a 450 test is worth 32 on a 400 test or 22 on a 500 test.

Types of question at the various levels

While all the levels have a strongly structural base, there are different types of content appropriate to the various levels. These include items for vocabulary, idiom and pronunciation. See Table 2 (page 15) for a breakdown of contents at the various levels.

Table 1: Relationship between proximate levels

A												
Score of	on Test	is equivalent on Test →	050 to↓	100 to↓	150 to↓	200 to↓	250 to↓	300 to↓	350 to↓	400 to↓	450 to↓	500 to↓
27	050		27	22	17*							
30	050		30	25	20*							
33	050		33	28	23*							
27	100		32	27	22	17*						
30	100		35	30	25	20*						
33	100		38	33	28	23*						
27	150		37*	32	27	22	17*					
30	150		40*	35	30	25	20*					
33	150		43*	38	33	28	23*					
27	200			37*	32	27	22	17*				
30	200			40*	35	30	25	20*				
33	200			43*	38	33	28	23*				
27	250				37*	32	27	22	17*			
30	250				40*	35	30	25	20*			
33	250				43*	38	33	28	23*			
27	300					37*	32	27	22	17*		
30	300					40*	35	30	25	20*		
33	300					43*	38	33	28	23*		
27	350						37*	32	27	22	17*	
30	350						40*	35	30	25	20*	
33	350						43*	38	33	28	23*	
27	400							37*	32	27	22	17*
30	400							40*	35	30	25	20*
33	400							43*	38	33	28	23*
27	450								37*	32	27	22
30	450								40*	35	30	25
33	450								43*	38	33	28
27	500									37*	32	27
30	500									40*	35	30
33	500									43*	38	33

* indicates that the relationship is not so accurate at two stages removed

Table 2: Composition of Tests: Types of Question used

	050	100	150	200	250	300	350	400	450	500
Structural: single item	50	50	38	36	24	29	0	0	0	0
Structural: continuous prose	0	0	12	14	13	8	33	33	25	24
Pronunciation: rhyming	0	0	0	0	6	6	5	5	0	0
Lexical	0	0	0	0	7	7	12	12	15	18
Idiom	0	0	0	0	0	0	0	0	5	4
Pronunciation: homonyms	0	0	0	0	0	0	0	0	5	4
Total	50	50	50	50	50	50	50	50	50	50

Key to Test 050 A

	A	B	C	D
1		B		
2				D
3		B		
4				D
5		B		
6	A			
7		B		
8				D
9			C	
10		B		
11			C	
12			C	
13				D
14		B		
15		B		
16				D
17	A			
18			C	
19			C	
20	A			
21		B		
22			C	
23			C	
24				D
25			C	

	A	B	C	D
26		B		
27				D
28				D
29			C	
30	A			
31		B		
32			C	
33	A			
34	A			
35		B		
36			C	
37		B		
38	A			
39		B		
40			C	
41		B		
42			C	
43				D
44				D
45			C	
46	A			
47				D
48	A			
49	A			
50	A			