

Guidelines for Writers of Reading and Use of English Tests

© BRITISH COUNCIL, HUNGARY

1 Introduction

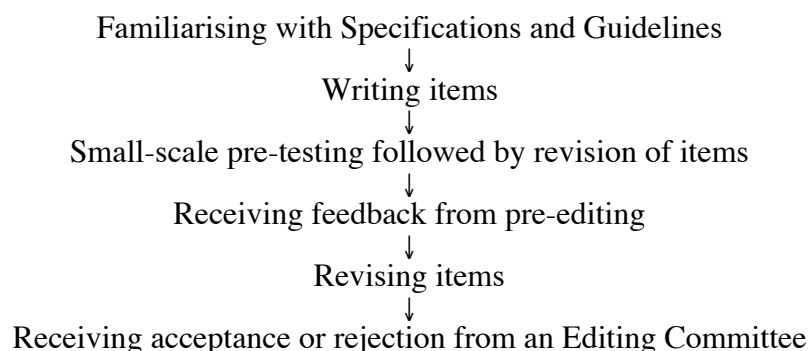
These Guidelines were originally written to help secondary school and university teachers, who had been trained in language testing, to write test items for the British Council English School-leaving Examination Reform Project. We believe that they are also useful for English teachers in general, even if they have not received specific training in language testing, to help them write better classroom language tests. And even if teachers do not wish or do not have to write their own tests, understanding how Reading and Use of English tests are constructed should help teachers understand the tests better, and should help them prepare their students better for modern European English examinations.

In this document, you will find both general and specific guidelines for the construction of Reading and Use of English test items and tasks. Each set of guidelines deals with text selection, task selection, task design, and rubrics.

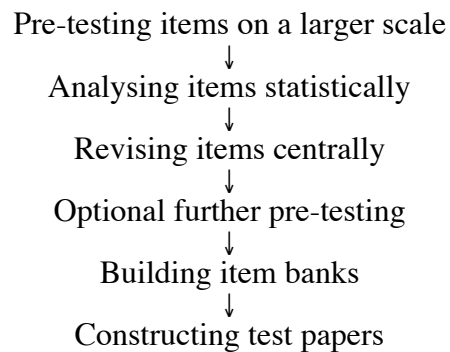
Normally, guidelines like these are accompanied by **test specifications**. Specifications are, if you like, the blueprint for what the test should contain. They indicate who the test-takers are (how old they are, what level of ability they have, and so on), what the purpose of the test is, the level of the test (for example in terms of the Common European Framework), what abilities the test is intended to test, the number of sections or separate papers in the test, the timing of each section or paper, the weighting of each section (how many points), which text types should be used, how long they should be, which language skills and elements should be tested, which test tasks or item types should be used, how rubrics should be worded, what criteria should be used for marking, and so on.

Test writers need to understand the specifications for the test they are writing items for, and they need to refer to them constantly. However, classroom teachers may not have any such specifications to refer to. Clearly test writing is easier if you have such specifications, and so it is often very helpful to develop specifications for tests you need to write, preferably in consultation with colleagues. Such specifications will help you write better tests. If you do not have, or do not know how to write, such specifications, do not despair! Simply following the Guidelines that follow will help you write better tests. And once you have written a test according to these Guidelines, it should become clearer to you what you need to decide in terms of what content, etc, your tests should contain. In other words, following these Guidelines will help you to see the need for specifications, and should also help you draft specifications for your own tests.

Item writing is part of the test construction process, and examination boards usually follow internationally accepted procedures. These procedures were also followed in producing the tasks in this book. The first part of the test construction process includes the following steps:



The second part of this process normally includes the following steps:



Such an elaborate process is often beyond the possibilities of classroom teachers, be they at secondary school or university. However, small-scale pre-testing on suitable students, perhaps in colleagues' classes, is often possible and is certainly highly desirable. And even if you do not have a formal Editing Committee, it is almost always possible to ask fellow English teachers to look over your test items, to check them for errors that you may not have seen, to check that they understand what they have to do, that they get the right answer, and so on. In fact, it is a general principle of language testing that it is impossible to write good test items alone. You ALWAYS need to get the opinions of colleagues about the items and wherever possible to try them out on students. Analysing test results after administering a test to your class is also a good way of seeing how well you wrote the test items. You can then revise and improve them for another occasion - you could even construct your own bank of test items for use again and again.

2 General guidelines

2.1 Text selection

- 2.1.1 All texts must be authentic.
- 2.1.2 Texts must not be taken from printed material in other examinations or English language teaching materials.
- 2.1.3 The topic of the texts must be accessible to the students' age group.
- 2.1.4 Texts must not be offensive, distressing or violent.
- 2.1.5 Texts must be of a suitable length.
- 2.1.6 Texts must be of a suitable level of difficulty.
- 2.1.7 A text must have a title unless it is used as part of the item.
- 2.1.8 The same text can be used with more than one task. For example, a reading text with different tasks could be used for intermediate and more advanced levels, or a reading text used with intermediate students could be used as part of a Use of English task for more advanced students.

2.2 Task selection

- 2.2.1 Have a clear idea what skill and sub-skill(s) you intend to test and select the task and text/prompts accordingly.
- 2.2.2 The task type must be one which is familiar to the students who are to be tested.
- 2.2.3 Tasks must be accessible to the students' age group.
- 2.2.4 Tasks must not be offensive, distressing or violent.
- 2.2.5 The time allocated for each task must be sufficient for somebody who has the ability being tested to complete the task comfortably within the time limit.
- 2.2.6 Students must be able to see easily how the task relates to the text.
- 2.2.7 No more than one testing technique should be used in each task. Do not switch from multiple-choice to single-word answers in a single task, for example.

2.3 Task design

- 2.3.1 There should be a minimum of five items in a task.
- 2.3.2 Items must follow the text sequence.
- 2.3.3 Items must be spread evenly through the text.
- 2.3.4 It must not be possible to answer any item without reference to the text. This must be checked carefully.
- 2.3.5 In non-sequencing tasks, items must not be interdependent –students should not need one answer in order to find another.
- 2.3.6 Items must not overlap.
- 2.3.7 Two items must not have similar answers.
- 2.3.8 Items must have a complete answer key. All possible good answers and predictable wrong answers should be provided.
- 2.3.9 The items in each task must be numbered sequentially.
- 2.3.10 One item in each task must have an answer provided as an example. The example item, written in italics, should be marked 0 and should precede the other items. If the options to be selected are marked with letters, the answer to the example should be marked with a letter according to its place among the options.
- 2.3.11 The example answer must be provided in the form in which students are expected to write their answers and, where appropriate, it must be hand-written.

- 2.3.12 “Find the wrong answer” type items are not acceptable in multiple-choice tasks.
- 2.3.13 Each item must score one point.

2.4 Rubrics

- 2.4.1 Rubrics should be in English. They must be clear, simple and brief. Redundancies must be deleted. Do not use exclamation marks and meta-language.
- 2.4.2 Conform to standard rubrics, for example, the ones in this textbook.
- 2.4.3 Indicate clearly what students have to do i.e. what they have to write and where.
- 2.4.4 The rubrics should be in bold.
- 2.4.5 Put the total number of points available for each task (in other words, the number of items) at the end of the task in bold.

2.5 Checklist

Text

- 2.5.1 Is the text type worth testing?
- 2.5.2 Is the text user-friendly in style and acceptable in content?
- 2.5.3 Is the length of text appropriate for your students?
- 2.5.4 Is the level of difficulty appropriate? Do any alterations to the text conform to the Guidelines?

Task

- 2.5.5 Is the rubric clear, brief and in bold?
- 2.5.6 Has an example been provided for each task?
- 2.5.7 Is there a complete answer key for each task?
- 2.5.8 Is the task type familiar to students?
- 2.5.9 Is the task accessible and acceptable?
- 2.5.10 Is the number of items acceptable?
- 2.5.11 Are the items spread evenly through the text?
- 2.5.12 Are all items properly labelled: questions with numbers, options with letters?
- 2.5.13 Can each question be answered independently?
- 2.5.14 Do items follow the text sequence?
- 2.5.15 Have you provided a full key?
- 2.5.16 Do answer keys reflect the form in which the answer is expected?
- 2.5.17 Does the answer key include alternative answers (if applicable)?

2.6 Acknowledgement

The following reference materials have been used in compiling these guidelines:

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge Language Teaching Library. Cambridge: Cambridge University Press.
- Milanovic, M. (Ed.). (1996). *User's Guide for Examiners*. Council of Europe, CC-LANG, 10.

3 Reading Paper

3.1 Text selection

- 3.1.1 Use texts that are authentic and suitable for the selected task type.
- 3.1.2 Select texts whose topic is accessible to the students' age group.
- 3.1.3 Try to select texts that are likely to be interesting for students. Be aware that humorous texts might be enjoyable and easy for you but very difficult for students.
- 3.1.4 Avoid offensive or distressing topics.
- 3.1.5 Avoid well-known novels as text sources.
- 3.1.6 Select texts of suitable length.
- 3.1.7 Do not use more illustrations than necessary. When illustrations are used, they should be clear and photocopiable.
- 3.1.8 Do not use published teaching materials.
- 3.1.9 Do not use texts with content covered in detail in any textbook. You do not want to test knowledge of content, but ability to understand texts.
- 3.1.10 Do not make any changes to the original text. Do not delete words, sentences or paragraphs from the selected body of text. If the text contains any offensive words that you think should be replaced, only change these with great care and always seek the advice of a fellow teacher or a native speaker as to the acceptability of the changes you have made.
- 3.1.11 If the text used is not the beginning part of a longer text, make sure it stands alone and does not contain references to previous parts of the original text.
- 3.1.12 Wherever possible, begin the text with its title or headline.
- 3.1.13 Edit the text layout so that it resembles the original text as closely as possible but make sure this does not make the text too difficult to read (e.g. in gap filling tasks).

3.2 Task selection

- 3.2.1 Have a clear idea about the purpose of each reading task, which reading skill is tested in a particular task, and why a particular task type is used.
- 3.2.2 The task should reflect the way in which people usually read the type of text used.
- 3.2.3 Consider applying more difficult texts with easier tasks and vice versa.

3.3 Task design

- 3.3.1 There should be a minimum of five items in a task.
- 3.3.2 Do not design in any one task more than 10 items for a passage of approximately 100 words.
- 3.3.3 The wording of the items should be such that it is easy to understand, is below the language level of the text and does not require more reading than the text itself.
- 3.3.4 Provide an example, written in italics and marked with (0), at the beginning of each task. Example answers should be given in a hand-written form where appropriate.
- 3.3.5 Make sure that no item can be answered correctly without reading the text.
- 3.3.6 Make sure that each question can be answered independently; that is, the answer to one question does not depend on the answer to another question (except in sequencing tasks).
- 3.3.7 Make sure that items do not overlap.
- 3.3.8 Make sure there is only one possible answer to each question. If more than one answer is required, this fact must be clearly indicated, and they must count as different items.

- 3.3.9 Avoid questions which demand much imagination.
- 3.3.10 Sequence items in the order they appear in the original text (other than sequencing tasks!).
- 3.3.11 Number items and use letters to mark options.
- 3.3.12 Avoid Yes/No questions and True/False items, because guessing the right answer is easier.
- 3.3.13 Do not delete the first words or beginnings of sentences.
- 3.3.14 In any gap-filling task (e.g. modified cloze, banked cloze, discourse cloze) there should be a sentence at the beginning (before the example) for lead-in and a sentence at the end for lead-out, which do not have any items.
- 3.3.15 In a gapped text, where you take out paragraphs, do not take out the first or the last paragraph, in order to leave enough context for comprehension.
- 3.3.16 In gap-filling tasks there should be at least five words between gaps to provide enough context.
- 3.3.17 In banked cloze tasks the items must be numbered in the gaps in brackets, and the options, marked with letters, must be listed in alphabetical order either at the bottom of the page or on the facing page.
- 3.3.18 Provide four options for multiple-choice items. Multiple-choice questions should have only one correct answer and should not offer “either a) or b)” or “both a) and b)” type of answers as an option, because these are tricky.
- 3.3.19 “Not mentioned” must not be used as an optional answer, because this can be misinterpreted. Many things “not mentioned” can still be inferred.
- 3.3.20 In one task, the number of matches must not be more than 10 for intermediate students, and 12 for advanced students. Provide more choices than matches in matching tasks and banked gap-filling tasks (minimum 1, maximum 2 distractors to be provided).
- 3.3.21 Make sure that ambiguous matches are excluded.
- 3.3.22 Make sure that in sequencing tasks there is only one correct order and no more than 6-7 items for intermediate students, or 8-9 items for advanced students. Be careful if using newspaper articles in sequencing tasks: they are rarely in chronological order.
- 3.3.23 Clearly indicate the required length of the answers (maximum 3 words) in open-ended (short-answer) questions.
- 3.3.24 Design tasks with clear and consistent layout. The input text should retain the original format as much as possible. Paragraph texts clearly with indentations.
- 3.3.25 Design one task (including rubric, text, questions, space or boxes for answers) to fit on one page. If this is not possible, the task must be arranged on two facing pages.
- 3.3.26 Provide answer keys reflecting the form in which answers are expected. Give all possible or acceptable answers. Unacceptable answers that are likely to occur should also be indicated. In tasks where the correct answer may be phrased in various ways, the answer key should provide content-related instructions concerning acceptable answers.
- 3.3.27 Weight items equally (1 item = 1 point).

3.4 Rubrics

- 3.4.1 Indicate what type of text students are going to read and/or the text source (preferably both).

List of reference books

- Alderson, J C (2000) *Assessing Reading*. Cambridge Language Assessment Series. Cambridge: Cambridge University Press.
- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge University Press.
- Alderson, J. C. & Urquhart, A. H. (Eds.) (1984). *Reading in a Foreign Language*. Longman.
- Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford University Press.
- Bachman, L. F. and Palmer, A. S. (1996) *Language Testing in Practice*. Oxford: Oxford University Press.
- Carrell, P. L., Devine, J. & Eskey, D. E. (Eds.) (1988). *Interactive Approaches to Second Language Reading*. Cambridge University Press.
- Harrison, A. (1983). *A Language Testing Handbook*. Macmillan.
- Heaton, J.B. (1998). *Writing English Language Tests*. Longman.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge University Press.
- Nuttall, C. (1996). *Teaching Reading Skills in a Foreign Language*. Heinemann Publishers. Second edition
- Urquhart, A. H. & Weir, C. (1998). *Reading in a Second Language: Process, Product and Practice*. Longman.
- Wallace, C. (1992). *Reading*. Oxford University Press.
- Weir, C. (1993). *Communicative Language Testing*. Prentice Hall.
- Weir, C. (1993). *Understanding and Developing Language Tests*. Prentice Hall.
- Practice test booklets and practice books for the Cambridge FCE and CAE examinations

4 Use of English Paper

4.1 Text selection

- 4.1.1 Choose relevant, interesting and intrinsically motivating topics for the age group of students concerned.
- 4.1.2 Use authentic texts, where possible. Texts taken from graded readers as they appear in the original are acceptable. Do not take texts or sentences from language teaching materials other than readers or from dictionaries.
- 4.1.3 Use texts at an appropriate level of difficulty. In order to ensure that the task tests Use of English and not Reading, select texts the language level of which is below the tested level.
- 4.1.4 Only make changes to the original text if absolutely essential. Do not delete words, sentences or paragraphs from the selected body of text. If the text contains any offensive words that should be replaced, or you feel that any other changes would be necessary, only make these changes with great care and always seek the advice of a fellow teacher or a native speaker as to the acceptability of the changes you have made.
- 4.1.5 Do not use texts exceeding 300 words.
- 4.1.6 Normally provide a title. It helps to contextualise the meaning.
- 4.1.7 Edit the text layout so that it resembles the original text as closely as possible but make sure this does not make the text too difficult to read (e.g. in gap filling tasks).

4.2 Task selection

- 4.2.1 The selected task type must be familiar to the students.
- 4.2.2 Do not use "Find the wrong answers" multiple-choice tasks, because these are very tricky and may test intelligence rather than language.

4.3 Task design

- 4.3.1 Test a range of different linguistic points in one task.
- 4.3.2 Balance the number of lexical and structural items in text-based tasks (except in word transformation tasks).
- 4.3.3 Use only one testing technique in one task.
- 4.3.4 There should be a minimum of ten, a maximum of 20 items in a task, in an approximate ratio of one item per 10-15 words in text-based tasks.
- 4.3.5 Highlight items clearly by using **boldface** (e.g. **1**. He (**live**) here since he was born.).
- 4.3.6 Provide an example in italics marked (0), and an example answer in a handwritten form. The example should follow the lead-in sentence.
- 4.3.7 Make sure that each question can be answered independently; that is, the answer to one question does not depend on the answer to another question (except in sequencing tasks).
- 4.3.8 Make sure that items do not overlap.
- 4.3.9 Number items and use letters to mark options.
- 4.3.10 Provide enough context in between gaps (at least five words) in gap-filling tasks.
- 4.3.11 In any gap-filling task (e.g. modified cloze, banked cloze, discourse cloze) provide a sentence with no items at the beginning (before the example) for lead-in and one at the end for lead-out. This will help students understand the context of the items.

- 4.3.12 In banked cloze tasks the items must be numbered in the gaps in brackets, and the options, marked with letters, must be listed in alphabetical order either at the bottom of the page or on the facing page.
- 4.3.13 When designing multiple-choice cloze tasks, consult good monolingual dictionaries and books that give synonyms and near synonyms.
- 4.3.14 Provide four options for multiple-choice items. Multiple-choice questions should have only one correct answer and should not offer “either a) or b)” or “both a) and b)” type of answers as an option, because these are tricky and difficult to understand.
- 4.3.15 Put only one item in a line in intrusive-word tasks, "spot-the-error" and "spot-and-correct" tasks.
- 4.3.16 Do not use an item at the beginning or at the end of a line in intrusive-word tests. Students should be required to write the word to be deleted in the space provided. Crossing out the intrusive word is not the right way to indicate the answer. Deleting an irrelevant word may make the student delete another word.
- 4.3.17 In intrusive-word tests the word to be deleted must not appear in the line more than once.
- 4.3.18 Design one task to fit on one page if possible, or on facing pages.
- 4.3.19 Design tasks with clear and consistent layout. The input text should retain the original format as much as possible. Paragraph texts clearly with indentations.
- 4.3.20 Provide enough space for answers.
- 4.3.21 Provide answer keys reflecting the form in which answers are expected. Give all possible or acceptable answers. Unacceptable answers that are likely to occur should also be indicated.
- 4.3.22 Each item scores one point.

4.4 Rubrics

- 4.4.1 Give the number of errors in "spot-the-error" tasks, and indicate that students must tick error-free lines. In "spot-the-error" tasks identified errors and identified error-free lines each score 1 point. In "spot-and-correct" tasks only corrected items score; no marks are given for errors identified but not corrected.

List of reference books

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge University Press.
- Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford University Press.
- Bachman, L. F. and Palmer, A. S. (1996) *Language Testing in Practice*. Oxford: Oxford University Press.
- Harrison, A. (1983). *A Language Testing Handbook*. Macmillan.
- Heaton, J. B. (Ed.). (1982). *Language Testing*. Modern English Publications.
- Heaton, J.B. (1998). *Writing English Language Tests*. Longman.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge University Press.
- Purpura, J. (2004) *Assessing Grammar*. Cambridge Language Assessment Series. Cambridge: Cambridge University Press.
- Weir, C. (1993). *Communicative Language Testing*. Prentice Hall.
- Weir, C. (1993). *Understanding and Developing Language Tests*. Prentice Hall.
- Practice test booklets and practice books for the Cambridge FCE and CAE examinations