

ESL Test Manual

Adopted by the ESL Examining Board during the meeting on 9 February 2015

Contents

Reader's guide	p. 3
Step-by-step plan for the test procedure	p. 4
Appendix 1: Further information on the quality aspects of tests	p. 9
Appendix 2: ESL Test Evaluation Form	p. 13
Appendix 3: Information on types of questions	p. 18
Appendix 4: Standardisation schedules for multiple-choice questions	p. 23

Reader's guide

This Test Manual starts off by giving a step-by-step plan for the entire procedure, from compiling tests to evaluating them. It gives a useful description of all the steps in the test procedure one by one. If you need a very brief and concise summary of the most important quality aspects involved in compiling tests, please go straight to Appendix 2. This appendix contains the ESL Test Evaluation Form with an overview of all factors the Assessment Committee takes into consideration when evaluating tests. Examiners can use this when compiling their own tests.

Appendix 1 takes a closer look at the content of the elements in the step-by-step plan and the criteria in the Test Evaluation Form. These are dealt with in the shape of general guidelines and the appropriate considerations. A number of crucial aspects that determine the quality of a test are discussed here, such as validity and reliability. Appendix 3 provides information on compiling various kinds of questions that can be included in tests. Appendix 4 deals with standardising multiple-choice questions and specifies how the frequently-used format comprising 40 questions with a choice of four answers should be standardised. In connection with this, we briefly touch on standardising tests with a choice of two or three answers to the questions.

This Test Manual fleshes out the requirements laid down in the Test Policy. Examiners should use this manual to justify their tests during discussions with their colleagues on the tests they devise and hold. The guidelines and considerations in this manual provide the Examining Board and the Test Committee(s) with useful tools for promoting and monitoring the quality of tests at ESL.¹

Apart from a few requirements taken from the Course and Examination Regulations, this Test Manual does not contain any specific and fixed 'requirements' that *have* to be complied with or 'rules' that *have* to be obeyed. However, there are a number of 'technical preconditions' – such as the minimum number of multiple-choice questions required to ensure that the test is reliable – which may only be departed from after consultation with the Test Committee or specialists at Risbo. The main point of this manual is to provide an orientation framework for examiners with test qualifications whose aim is to compile the best possible tests in collaboration with their colleagues.

¹ This Test Manual does not specifically deal with theses as a type of test. Although the same general guidelines do apply to this type of test, particular attention should be given to them in all programmes in connection with the thesis policy laid down in the programmes.

STEP-BY-STEP PLAN FOR THE TEST PROCEDURE

Step 1. Completing the test matrix

A test matrix is a useful tool for creating a framework for the composition of a test. A test matrix can be compared to a general table showing the course objectives to be tested, the cognitive (process) level that students must have, and the way in which the questions are distributed over the relevant subject matter and levels. Completing a test matrix will also enable you to compile more tests with a comparable content (such as resit exams).

Clearly-formulated course objectives are essential (use action verbs in the course objectives and formulate the course objectives in the positive):

- Re-read the course objectives.
- Establish the level at which they are formulated: B1 – B2 – B3 – Master? In this connection, establish whether the most important factor here is memorising, understanding, applying, analysing, evaluating or creating.²
- Check that the content and level correspond with the final attainment level and the Test Plan;
- Check whether it is possible to measure study progress in these objectives;
- Determine the relative weight of the objective. How many points must be assigned to the appropriate questions in the examination?

For example, a test matrix (partially completed) for a test at B3 or Master level would look like this:

Course objective/ subject	Cognitive (process) level						Total points
	Memorising	Understanding	Applying	Analysing	Evaluating	Creating	
Course objective A	10 points						10 points
Course objective B			15 points				15 points
Course objective C		20 points					20 points
Course objective D		20 points					20 points
Course objective E				20 points			20 points
Course objective F					15 points		15 points
Total	10 points	40 points	15 points	20 points	15 points		100 points

Step 2. Select the type of test(s) and the type of question(s)

Select a certain type of question or a combination of types of test questions by taking the following into consideration:

- the levels of the course objectives
- the available construction time required
- the available evaluation time required
- the time available for holding the test

² For more background information on these levels, please see Krathwohl (2002). A revision of Bloom's taxonomy. *Theory into practice*, 41 (4), pp. 212-218.

- familiarity with the different types of tests

When completing your test matrix, you can state in each cell what type(s) of questions will be used in the relevant test and the number of the questions used. It is quite possible that different parts of the test will have their own specific form.

The length of the test cannot be specified unequivocally, but the number of points to be obtained and the number of questions must be linked in such a way as to enable sufficient differentiation between the students, and that the measurement is reliable. You should also determine the ratio between relatively simple questions and relatively difficult ones (e.g. 60%-40%).

The table below gives all the answer times required (for closed-book examinations, i.e. excluding the consultation of collective legislation etc) for the most frequently-occurring types of questions:

<u>Type of question</u>	<u>Time required³</u>
correct/incorrect question	50 seconds
multiple-choice question with 2 answer options	50 seconds
multiple-choice question with 3 answer options	60 seconds
multiple-choice question with 4 or 5 answer options ⁴	60 seconds
open question (one-word or one-sentence answer)	1 minute
open question (answer takes up a quarter of a sheet of A4)	5 minutes
open question (answer takes up half a sheet of A4)	10 minutes
open question (answer takes up one sheet of A4)	25 minutes
open question (answer takes up two sheets of A4)	60 minutes

See Appendix 3 for more detailed information on the different question types.

Step 3. Compiling the test

In this phase, you will actually be writing down the questions. Devising the questions should preferably be done at the same time as the development of the teaching material. It is a good idea to devise more questions than you need, and also to devise extra questions for resits. If you do this, you can delete all the questions that are less satisfactory when going through the test again. You can save the best questions for other examinations later on.

When devising your tests and test questions, you should use the ESL Test Evaluation Form as a checklist (see Appendix 2).

Step 4. Writing down an answer model and score instructions; determining the marks to be awarded

You can base an answer model (for open questions) on your own preferences, as well as on the answers your colleague gives to the questions you have compiled. An answer model enables you to increase a test's objectivity and therefore its reliability.

³ The time stated for the multiple-choice questions concerns closed-book examinations in which you expect students to be able to indicate the correct answer very quickly. Students will need more time for open-book examinations and in cases where multiple-choice questions are accompanied by case histories.

⁴ A test that *entirely* consists of four-choice questions must comprise at least 40 questions. For three-choice questions, the number of questions in the test must be 60 and for two-choice questions, the number of questions must be 80.

You use the answer model to compile and determine score instructions (also known as an assessment model). This means that you enter the maximum number of points that each student can obtain if they answer questions or parts of questions correctly in your answer model. You can consult the test matrix you have compiled to obtain the total number of points for the question and the entire test. Set assessment criteria in consultation with your colleagues.

Step 5. Asking a colleague to do the test and evaluating it afterwards

Ask one of your fellow lecturers to do the test. This should preferably be someone in the same department as you, so that any dubious questions and answers can be deleted. You could also ask your colleague to consider the following: is the question relevant to the objective, and is it important for the student to be familiar with what is being asked or able to do it? Help from your colleagues has a particularly favourable effect on test reliability. Asking a colleague who is not familiar with the course to do the test is also very instructive. For example, this can help you estimate the time required (please note: if you allow 2 hours, this usually turns out to be 2.5 to 3 hours in the actual examination) and the probability of students guessing the answers. In addition, you can use the Test Evaluation Form to evaluate the test.

Step 6. Writing down the instructions for the students

You should give proper instructions for completing the test in order to ensure unequivocal answers. Another extremely important point is transparent communication about the course and the associated test. In this respect, you should differentiate between notifications at the start of the course and instructions on the test cover page.

Notifications at the start of the course (please include these in the course information under the heading 'Testing')

- when will the test be held (date, time, duration)?
- where will the test be held?
- do students have to enrol for the test, or can anyone join in if they like?
- what rules apply to absence for a test?
- are resits available?
- what form will the test be in?
- how many questions can students expect, and what types of questions will be asked? (Give the students a practice exam paper or practice questions.)
- what criteria does students' performance have to comply with?
- what do the students have to take along to the test?
- what materials and tools will be provided by the lecturer?
- how will students be notified of the test results?

Information on the test cover page:

- how many questions will there be in the examination, and what types of questions will be asked?
- how much time will the students have?
- what are students allowed to use?
- instructions for completing the test
- how will the test be marked? how will it be evaluated?
- what are the consequences of breaking the rules? how do sanctions work?

N.B.: A 'standard cover page' for examinations has been developed thanks to the Test Committee. You can obtain a copy of this from the secretariat of the TC/Examining Board. The standard cover page includes standard elements as well as variable choice elements.

Step 7. Holding the test

This step is self-evident: the students will do the test. An examiner will be present during the first hour of the test at the very least. After this, the examiner will be available by telephone to answer any questions the students may have.

Step 8. Posting the key/model answers

You should give the students the correct answers to multiple-choice questions immediately after the test. Invite the students to send any concerns to the course coordinator.

Step 9. Checking the answers

Check the answers to the open questions using the answer model/assessment model as a basis. For open questions, work with at least two other evaluators working independently of each other if possible. In the case of a small number of students whom the lecturer knows well, we recommend that you work with figure codes on the answer sheets instead of names. Check the work preferably for each individual question, not for each person. First read a number of answers for each question; this will give you an idea of the feasibility of the standard to be applied. Change the order of the pile of answer sheets now and again. If possible, check the work without taking long breaks; this will ensure that your assessments will be as homogeneous as possible.

Standardisation/awarding marks for multiple-choice questions

When working with multiple-choice questions, you should bear in mind that students may have a guess at the answers. For multiple-choice examinations with four answer options, this is one-quarter of the total number of questions, one-third for three answer options, and half for two answer options. This will be taken into consideration during the automatic awarding of marks when processing the examinations. For further details, please see Appendix 4: 'Standardisation of 40 four-choice questions'.

N.B.: After assessing the answers, you sometimes need to adjust the standardisation. This will result in the marks awarded being different to those you would have obtained from the previous standardisation you set. In such cases, you should apply the Protocol laid down by the Examining Board.

Step 10. Analysing the test results

Studying the test results

The next step concerns an analysis of the quality of the test as a whole and of each separate question in the test, on the basis of studying patterns in the answers given by the students. Naturally, the most important points here are validity and reliability. A comparison of the test questions with the answers given by the students will provide important material. If examinations containing multiple-choice questions are processed automatically, the system can turn these into statistics. This has to be done manually in the case of open questions.

It is important to check carefully whether any mistakes have been made in the marks awarded or whether anything in the questions is not absolutely clear, or whether, on further inspection, it emerges that several answers are possible, or whether the questions differentiate positively between the students, and so on.

The quality of an examination is logically determined by the sum of its separate parts. Therefore, if there are any problems with the quality of the individual questions, you could consider rectifying this by removing one of the questions from the examination, or in some cases, by marking a second alternative as correct. However, this will affect the reliability of the measurement as well as the representativeness of the test, so you should think very carefully before doing it and you should certainly not do it too often.

Guideline: Your total aim should be to mark two alternatives as correct for a maximum of 10% of your questions, and to delete or compensate for a maximum of 10% of the questions. If you depart from this, or if you want to change the standardisation, you should apply the Protocol for departing from or adjusting your own assessment standards by sections, in consultation with the Test Committee.

Examine students' criticism of the questions

Students can make a major contribution towards assessing the quality of the test questions. So it is a good idea to make a point of asking your students to share any comments or complaints they may have with you after the examination. Students can point out anything that is unclear in the questions or anything that could be misinterpreted, which you yourself might have overlooked. Please make sure that you inform the students what use you have made of their comments and remarks in the explanatory notes to the test results no later than the date on which the marks are posted.

Step 11. Determining the marks and posting them

Determine the marks after considering the points stated in Step 10.

Post partial scores and grades that are not rounded off on LawWeb. Students will appreciate it if they are able to inspect the assessment standard used at the same time. Have a think about whether it would be useful to post partial grades, e.g. on various open questions or the results of the multiple-choice part of the test. Let the students know where, when and how the follow-up meeting or inspection will be held.

If, after consultation with the Test Committee, the standardisation has been adjusted to a substantial extent, you must inform the students on the differences between the standardisation you ultimately used and the information you gave them about this beforehand. It is very important that you explain this change clearly to them.

It is equally important that the only students who are allowed to inspect their test results are the ones who were actually entitled to do the test. Students who were not entitled to do the test should be allowed inspection of their results through the Examining Board if the Board decides that the marks obtained really will be awarded.

Step 12. Discuss the results with the students and give them the opportunity to inspect their test papers

This final phase, i.e. the follow-up meeting, is extremely important to students' learning process and to the transparency of the entire test procedure. Feedback, preferably as specific as possible, can be given individually or in a group. It is very important that this is done within the period of time set beforehand. Please make sure that all students who are unable to attend the follow-up meeting are still given the opportunity to inspect their test paper. Provide them with the answers given at their request, possibly for a small fee: say € 1 or € 2.

Make it clear to them that no discussions on specific assessments will be held during the follow-up meeting or inspection. Any requests for adjustments to be made to the assessment must be submitted to the course coordinator in writing, stating the reasons for the request and

including a reference to the assessment standard. The test should preferably be re-assessed by a second assessor. It is a good idea to keep a 'log book' to record the results of any re-assessments for each student. Inform the relevant students through a general announcement that a reassessment has been carried out and if relevant, that their marks have been adjusted accordingly.

Appendix 1 Further information on the quality aspects of tests

1) *Validity of the test*

Guidelines:

- The test must be compiled in relation to the programme's Test Plan in which the final qualifications for the programme are linked to the specific objectives for the course in which the student is being tested.⁵
- Create a test matrix with a diagram to make clear what the relationship is between the course objectives, the distribution of points and the different types of questions or tests in the Test.
- Select the types of test that fit in best with the course objectives.
- As far as possible, set questions that are consistent with the concrete elaboration of the course material. If possible, make sure you have finished compiling the test before the start of the term.
- Compile a lot more questions than the test matrix actually requires. Bear in mind that not all questions will 'survive' your colleagues' evaluation. In addition, try and make sure that questions for the examination and the resit are compiled at the same time.
- Compile your questions and the test in collaboration with all your colleagues who are directly involved in the course.
- Ask your colleagues to assess the draft of your test.

Considerations:

- The validity of the test indicates whether the test measures what it is supposed to measure. To this end, you should carefully consider what types of tests fit in the best with all the different course objectives. It is essential to ensure satisfactory coordination between the course objectives, skills, and the type and content of the test.
- Creating a test matrix requires your explicit consideration of the level of mastery expected of the students in respect of the course objectives. The level of difficulty in the test must correspond to the phase of the programme in which the course is given. The matrix is also a useful tool for compiling a balanced test whereby all the course objectives and the relevant important points are included in the test. The test matrix may also be useful as a benchmark for subsequent tests.
- A good test requires clearly-formulated course objectives. General and abstract phrasing is not at all practical; this should be further elaborated in concrete terms at the very least. After this, you should check that the type of test fits in with the course objectives.
- And finally, you should check that no subjects are tested unless these have been announced beforehand. You must inform the students of the course objectives, types of test and general assessment criteria well in advance.
- Students are entitled to an example of an examination paper. You can use a previous examination paper for this, but you should make sure that this previous examination fits in well with the present course.

2) *Representativeness of the test and test questions*

Guidelines:

- Use your test matrix to ensure that the questions and the points allocated for each question are evenly distributed across the course objectives and the subject matter.
- Make sure that the distribution of questions on knowledge, insight and application fits in well with the level at which the course is given, in which respect the students are expected to have sufficient mastery.
- Determine the number of questions on the basis of the number of matters to be addressed.

⁵ The ESL Test Policy states that a Test Plan must be drawn up for each programme. See the relevant document.

- At least 55% of the questions must relate to the compulsory written subject matter dealt with during tutorials. A limited number of questions (no more than 25%) may relate to subject matter exclusively addressed during lectures; essential information should preferably be traceable on the sheets. All questions or assignments in the examination relating to Practicals may determine no more than 20% of the final marks.

Considerations:

- Is the test representative of the subject matter that the test is supposed to be examining? Is the test unbiased? Have the questions and assignments been sufficiently weighed up against each other in respect of the knowledge, insight and skills to be tested (and possibly students' approach to their studies)? The questions and assignments must relate to the course objectives formulated, and the questions must be evenly distributed across all the different course objectives. In this case, too, the test matrix gives you a useful overview of the distribution and level of difficulty of the questions.

3) Reliability of the test and test questions

Guidelines:

- Make sure that all the questions and assignments in the tests can only be interpreted in one way.
- Although more questions increase the test's reliability, they also require more time. Ensure a good balance and include sufficient different questions to obtain sufficient insight into the student's mastery of all the course objectives.
- Make sure you phrase the assignments in the tests as briefly, concisely and clearly as possible.
- Make sure that the students are able to do the test with satisfactory results in the time allocated for the test.
- Always determine the definitive shape and content of a test in consultation with one or more of your colleagues, or after such consultation.
- In consultation with those of your colleagues who are directly involved, set criteria for assessment and incorporate these criteria into an assessment model.
- At the start of the assessment, ask your colleagues to check whether this model needs to be adjusted. To this end, ask all the assessors to evaluate the same random sample of answers and discuss the differences in the allocation of points in order to calibrate the application of the model.
- Reliability among assessors: if a number of your colleagues are involved in the assessment of the students' answers, it is essential to guarantee parity in the assessment.

Considerations:

- Reliability of a test should be taken to mean the extent to which the measurement (the score) reflects the student's actual level of mastery, and is unencumbered by unintentional and coincidental influences. Longer and more comprehensive tests are more reliable than shorter ones. The main principle is that the more valid and representative questions there are in a test, the more reliable the test as a whole, provided that the available time is taken into consideration.
- The reliability of a test is also greater if various assessors arrive at the same allocation of points for the same answers on the basis of the assessment model.
- For the sake of transparency, make sure that the assessment model – or a version of it – is made available to the students afterwards.

4) Specificity of the test and test questions

Guideline:

- There must be a demonstrable difference in the answers to the questions given by students who have seriously studied for the test and those given by students who have not prepared sufficiently for it.

Considerations:

- In some test situations, participants can get quite a long way by using their common sense. In that case, there is something wrong. The more a student relies on their common sense, the more we may assume that the distribution of the scores will be slight and that the test will lose its value and reliability.
- The specificity of a test can be established by asking persons who have followed the curriculum to do the test, as well as persons who have not followed the curriculum. If the test is sufficiently specific, the latter group will not be able to do the test successfully compared to the first group.

6) Transparency on the nature, content and assessment of the test beforehand

Guidelines:

- Notify the students of the course objectives in the information on the course, and make sure the test fits in well with the priorities set in this respect during the course.
- Inform students of the type and length of the test, the nature of the questions, and the amount of time they will have to complete the test.
- Ensure that the students can form a concrete idea of the test and allow them to practise using representative sample questions. For example, you could hand out one of last year's examination papers during the course, or hold a practice exam.
- Tell the students what factors they will be assessed on and what standard is used in this respect, and give them a rough idea of how the points are distributed across the different parts of the test.
- Determine a time period within which feedback will be given.

Considerations:

- It is essential that the emphasis on course objectives in education fits in well with the choice of type(s) of test and the content and level of difficulty of the parts of the test, and that the students are kept well informed on these matters.
- In order to be able to prepare properly, students have to know what to expect in the tests and all the different parts. Students should be clearly informed on the subject matter to be assessed and how this will be done, the level of mastery required, what they are and are not allowed to do during examinations, the consequences of any irregularities, etc.
- You could help students know what to expect by e.g. handing out an examination paper from the previous year. If you do this, students will not be taken unawares by certain types of questions; they will know how much time they will be given for answering the questions and they will be able to familiarise themselves with the way in which the test is set up and held.

7) Transparency on the nature, content and assessment of the test afterwards

Guideline:

- Make sure you give the students sufficient feedback by providing them with a clear written explanation of the way in which the assessment was carried out (use the assessment model to do this).
- If all or part of an examination comprises multiple-choice questions, make sure that you give the students the right answers as soon as possible after the examination. When doing

this, invite the students to draw the examiners' attention to any critical comments the students may have. Get them to do this as soon as possible.

- Organise a collective follow-up meeting during which the assessment (assessment model) is explained and students can ask further explanatory questions.
- Create a procedure in which students are able to respond individually to the way in which their answers have been assessed, and in which they can request a re-assessment of their answers, thereby stating their reasons.
- Make it clear to the students that in the case of a re-assessment, all the test results will be scrutinised.
- Ensure that the students are able to inspect their answers after the follow-up meeting as well. For example, you could organise this by providing students with a copy of their answers on request, for a reasonable fee.

Considerations:

- Transparency after the examination means that students should be able to understand how their answers have been specifically assessed. To this end, they must have insight into the method and the criteria used during assessment. They must also be given access to their answers in order to judge whether they have actually understood why their answers have been assessed in that way. If this is not the case, they must be given the opportunity to ask for a further explanation in this respect. In this connection, any student who is not convinced in the first instance must also be given the opportunity to request a general re-assessment, thereby stating their reasons.

Appendix 2 ESL Test Evaluation Form

A. Evaluation criteria at test level	Does the test fulfil the criterion?	Explanation of points for improvement (if appropriate)
<i>Validity</i>		
The learning objectives to be tested were made clear when doing the test.	Satisfactory/unsatisfactory	
The learning objectives are present in the test.	Satisfactory/unsatisfactory	
The questions and assignments in the test do not exceed the examination subject matter stated.	Satisfactory/unsatisfactory	
No more than 25% of the test questions relate to examination subject matter exclusively dealt with during lectures ⁶ .	Satisfactory/unsatisfactory	
The type of test and the type of questions fit in with the cognitive level of the learning objectives and the course methods.	Satisfactory/unsatisfactory	
The number of test questions for each learning objective and the number of points that can be obtained in this respect fit in with the level of difficulty of the relevant learning objective.	Satisfactory/unsatisfactory	
<i>Reliability</i>		
The questions in the test are formulated independently of one another (on the one hand, an error in Question A will not result in errors in Questions B or C, while on the other hand, Question X does not reveal any information that is useful for answering Question Y).	Satisfactory/unsatisfactory	

⁶ Taken from the ESL Course & Examination Regulations.

A. Evaluation criteria at test level	Does the test fulfil the criterion?	Explanation of points for improvement (if appropriate)
The test comprises sufficient questions to ensure a reliable measurement. The guideline for multiple-choice questions (also in combination with open questions) is as follows: there must be at least 80 questions for two-choice answers, at least 60 questions for three-choice answers, and at least 40 questions for four-choice answers.	Satisfactory/unsatisfactory	
The length of the test is proportional to the duration of the test, so that the students are able to answer the questions within the time specified, in accordance with reasonable standards.	Satisfactory/unsatisfactory	
<i>Transparency</i>		
The layout of the test and the order and grouping of the questions are logical. The guideline in this respect is that questions are grouped according to each subject, and each subject is briefly introduced.	Satisfactory/unsatisfactory	
The layout of the test is clear.	Satisfactory/unsatisfactory	
The numbering of the questions is logical and easy to follow.	Satisfactory/unsatisfactory	
There is a key or correction model, which is available to all assessors and which has been coordinated among all the relevant lecturers.	Satisfactory/unsatisfactory	
The answers in the correction model are clear and can only be interpreted in one way.	Satisfactory/unsatisfactory	
If applicable, each question makes it clear what components the answer must comprise and the number of points to be allocated for each component.	Satisfactory/unsatisfactory	
The examination contains clear test instructions for students and invigilators alike.	Satisfactory/unsatisfactory	

A. Evaluation criteria at test level	Does the test fulfil the criterion?	Explanation of points for improvement (if appropriate)
The learning objectives to be tested have been clearly communicated to the students.	Satisfactory/unsatisfactory	
The number of points to be allocated for each question or part of a question has been clearly communicated to the students.	Satisfactory/unsatisfactory	
The pass-fail border has been clearly communicated to the students.	Satisfactory/unsatisfactory	

B. Evaluation criteria at question level (multiple choice)	Note down the questions that fulfil this criterion	Note down the questions that do <u>not</u> fulfil this criterion
<i>The questions</i>		
The questions are grammatically correct.		
The language used in the questions is clear, concise and unambiguous.		
Only one subject or problem is addressed in each question.		
The question is asked in the positive (i.e. not in the negative unless absolutely necessary).		
Negatives such as <u>no</u> or <u>none</u> are underlined and there are no unnecessary double negatives in the questions.		
The questions are <u>not</u> trick questions.		
The questions do <u>not</u> contain any unintentional tips for answering.		
If possible, the questions are phrased in such a way that students are able to answer them without having to study alternative answers first.		
All introductory information is clearly separated from the questions to be answered or assignments		

B. Evaluation criteria at question level (multiple choice)	Note down the questions that fulfil this criterion	Note down the questions that do <u>not</u> fulfil this criterion
to be carried out.		
<i>Alternative answers</i>		
The content of the alternatives is equally formulated. The guideline here is that no alternative answers, such as 'none of the above alternatives is correct' or 'all the above alternatives are correct', are given under the question.		
The alternatives are phrased in the positive wherever possible. Any <u>negatives</u> are underlined.		
The length, level of detail and format of the alternatives are as similar as possible.		
The question and the alternatives fit in grammatically with one another.		
There is no overlap between the alternatives. The guideline here is that the 'incorrect' alternatives must be 100% incorrect.		
The alternatives are given in a logical order (or in a consistently random order), such as alphabetical order or numerical order.		
There are no repetitions of a major term used in the question in any of the alternatives given ('distractors').		
The language used in the alternatives is clear, concise and unambiguous.		

C. Evaluation criteria at question level (open questions)	Note down the questions that fulfil this criterion	Note down the questions that do <u>not</u> fulfil this criterion
--	---	---

C. Evaluation criteria at question level (open questions)	Note down the questions that fulfil this criterion	Note down the questions that do <u>not</u> fulfil this criterion
<i>The questions</i>		
The questions are grammatically correct.		
The language used in the questions is clear, concise and unambiguous.		
Only one subject or problem is addressed in each question.		
The question is asked in the positive (i.e. not in the negative unless absolutely necessary).		
Negatives such as <u>no</u> or <u>none</u> are underlined and there are no unnecessary double negatives in the questions.		
The questions are <u>not</u> trick questions.		
The questions do not contain any unintentional tips for answering.		
The questions give sufficient information on the length and form of the answers, such as the maximum number of words or specific features or aspects to be discussed.		
All introductory information is clearly separated from the questions to be answered or assignments to be carried out.		
If an answer has to be explained, substantiated or justified, this is clearly stated (possibly including the number of arguments to be put forward or the length of the explanation).		
The questions make it clear what content – i.e. knowledge, insight, skills, etc – is being tested.		
The questions test the use of objective knowledge and/or insight; they do not ask students to give		

C. Evaluation criteria at question level (open questions)	Note down the questions that fulfil this criterion	Note down the questions that do <u>not</u> fulfil this criterion
their own opinions, unless students are explicitly supposed to give their opinions and substantiate them with arguments.		
<i>Context information in test questions</i>		
Use of context – drawings, diagrams, texts, images, etc – is functional.		
The references made in the questions to texts, drawings and so on are correct.		
The context is clear and correctly reproduced.		
The context does not contain any irrelevant information.		

Appendix 3 INFORMATION ON TYPES OF QUESTIONS

We can distinguish the following types of questions:

1. multiple-choice questions
2. true-false questions
3. questions where students have to fill in the blank spaces
4. multiple-answer questions
5. matching questions
6. ordering questions
7. open questions

There is a great deal that can be said about the classic multiple-choice questions in particular; the information contained in these questions often relates to the other types of questions as well.

Re 1. Multiple-choice questions

General

- Tests must comprise a large number of questions, since a greater distribution across the total subject matter results in greater representativeness. The more questions in a test, the more reliable the test. The desired length of the test is related to the aim of the test and its duration.
- In the first instance, you should compile 50% more questions than the number that is ultimately included in the test, since a number of questions are always deleted on further inspection.
- With respect to the desired length of a test in a test for the final assessment, the test literature gives us the following guidelines:
 - at least 40 questions for four-choice questions
 - at least 60 questions for three-choice questions
 - at least 80 questions for two-choice questionsIt is advisable to submit cases that depart from this in the draft examination to the Risbo experts and ask their advice.
- Ask those of your colleagues who are experts in content to have a look at your questions. It is essential to check for deficiencies relating to form and content before the test is held in order to avoid any problems afterwards. You could do this during a group meeting of lecturers, for example. In this respect, Step 1 would be to ask the lecturers to answer the questions themselves, Step 2 would be to compare the answers you get (if there are any differences, you can improve the relevant questions and/or improve the 'stem' or the alternatives), and Step 3 would be to make all the other improvements needed.

Basic guidelines for phrasing

- During final tests, we recommend a type of question in which one of the alternatives provides the correct answer, rather than a type of question containing propositions that the students have to assess. This frequently-used type of question containing two propositions where students have to enter 'correct/incorrect' often gives rise to discussion. For example, students might quite justifiably ask themselves why they are not awarded any points for the right answer to half the question.
- Multiple-choice questions preferably contain three alternatives: the key and two 'distractors'. According to research, this is the best way to ensure efficiency, reliability and

speedy construction. Incidentally, four-choice questions also occur frequently because the probability of getting the answer right by guesswork is smaller here. However, the fourth alternative is often less effective as a 'distractor'.

- The alternatives for each question should be listed in a logical order, e.g. alphabetical order for words and numerical order (from 1 upwards) for figures. This will make the alternatives easier to follow while reducing the probability that students will seek to 'manipulate' the test constructor.
- Each question must state the final attainment level or course objective to which the question relates. Or in other words: you should avoid including irrelevant questions.
- Questions must focus on one sole idea, fact or problem.
- In addition, questions may not contain any grammatical or other indications for the correct answer. For example: 'One of the animal species that has become extinct in the Netherlands is ... (a) the badger (b) the robin (c) the white mountain rabbit'.
- Multiple-choice questions may not ask students to give any *general* opinions. In questions requiring an opinion, always give the context, such as the name of a philosopher.
- The 'stem' in a multiple-choice question must contain a problem or a question. It should not be so neutral that the question is actually contained in the alternatives instead of the 'stem'.
- Alternatives must all be more or less the same length (the correct answer is often too long compared to the 'distractors'!)
- Avoid word repetitions in the alternatives; it is better to include any repetitions in the 'stem'.
- In the case of 'supplementary questions', it must be possible to place the alternatives at the end of the 'stem'.
- All your alternatives must be plausible, as an alternative that can immediately be rejected increases the probability of guesswork.
- Questions may not contain any hints on which is the correct alternative.
- Questions must be brief and concise, without any superfluous information.
- Questions should preferably not be phrased in the negative. If this is unavoidable, you should draw attention to the negative phrasing by bold type or underlining.
- Avoid the following:
 - double negatives
 - vague indications such as 'sometimes' or 'many'
 - absolute contentions such as 'never' or 'always'
 - 'none of the alternatives stated are correct'
 - starting a question with the missing word
- Always use language that fits in with the subject matter studied
- All alternatives should automatically exclude one another.

Possible variations

- Direct questions: 'What date does the summer begin on?' a, b, c
- Supplementary questions: 'Summer begins on ...' a, b, c
- '- *Case* - Which of the rules of law stated are of importance in the case given?'
- 'Imagine that - *description* - What concept applies to this?'
- 'A and B are married under a pre-nuptial agreement (total separation of property). What is A not allowed to do without B's consent?'
- Placing concepts in the right order.

Re 2. True-false questions

General

- The idea behind this frequently-used type of question is that all knowledge can be expressed in propositions that are either true or false. The more familiar the students are with the subject matter, the better they will be able to identify the true and false propositions in this subject matter.
- All propositions must be 100% true or 100% false. In the case of subject matter where different authors hold different opinions on the same matter, there is a risk that statements will not turn out to be 100% true or false.
- Please note: if students make a guess, they can give the right answer to 50% of the questions. Of course, you can correct this when determining their final marks, but it would still be a pity if none of these questions were to count. In that case they would merely serve to ‘flesh out’ the test. This is also a drawback in diagnostic tests (without marks), as there is a considerable probability that students will guess the right answer and consequently think that they have a good mastery of the subject matter.

Basic guidelines for phrasing

- Make sure you phrase questions efficiently.
- Questions should preferably be phrased in the positive.
- Make sure your questions are unequivocally true or false.
- Make sure that propositions do not contain any half-truths or partial truths.
- Don’t use any words that give a clue to the answer.
- Explicitly draw students’ attention to propositions in the negative, such as by underlining them.
- Give approximately the same number of true and false propositions.
- Make sure that there is no ‘answer pattern’ in the order of the propositions.
- Avoid using questions taken straight out of the books studied.
- See also the basic guidelines for multiple-choice questions.

Possible variations

- ‘Bare-bones’ propositions: ‘Law of persons and family law form part of public law’
- Contexts containing one proposition: ‘Jaap sells his house to his sister Rita for 300,000 Euros. This means there is a property law contract between Jaap and Rita’
- Contexts containing a number of propositions: ‘Rita is divorced and wants to sell the house she is living in. Her brother Jaap buys this house at a fairly low price of 50,000 Euros. After a few months, Rita feels she isn’t so happy with this sale after all. But Jaap has already sold the house to Peter. State whether the propositions below are true or false. A property law contract exists between:
 - Rita and Jaap
 - Jaap and Peter
 - Rita and Peter

Re 3. Questions where students have to fill in the blank spaces

General

- This type of question is in fact an ‘open question type’: students have to formulate the answers themselves. It is difficult to use this type of question for digital tests because the marks cannot be unequivocally awarded (e.g. capital letters, typing errors, abbreviations).

Basic guidelines for phrasing

- Phrase certain statements, propositions and so on and leave out any key words or important parts of a sentence. Keep to the bare essentials.
- Phrase your questions in such a way that the blank space is at the end of each question.
- Phrase your questions in such a way that there are no more than two blank spaces. More than two blank spaces will make the test look like a crossword puzzle.
- Make sure that the questions contain all the necessary information to enable the students to answer them properly.
- Phrase your questions in such a way that only one correct answer is possible.
- Ensure that the answer space is the same size for each question, otherwise the students will draw conclusions from them.
- Make sure that the questions are not taken from sentences in the literature studied.

Re 4. Multiple-answer questions

General

- This type of question provides various answer options and asks students to place a cross next to the correct answers to each proposition.
- When compiling this type of question, you should have a good think about the marks. Does each correct answer selected count, or does only a completely correct answer count?
- See also all recommendations for multiple-choice questions.

Re 5. Matching questions

General

- Matching questions comprise two columns containing data. The data in both columns that belongs together must be combined in all cases. You can include different questions and different numbers of answers in each question. You can also arrange for points to be awarded on matching questions that have been partially correctly answered.
- The drawback to this type of question is that one single error can result in several other errors. For this reason, you should include more elements in the first column than in the right-hand one. Since it is quite difficult to ensure unambiguity when constructing this type of question, it is not much used in practice. This type of question is mainly used to test students' knowledge of facts.

Re 6. Ordering questions

General

- This type of question gives a number of answers that students have to place in the right order, e.g. chronological order. You can also arrange for points to be awarded on questions that have been partially correctly answered. This type of question is not much used in practice either. It is mainly used to test students' knowledge of facts.

Re 7. Open questions

Basic guidelines for phrasing

- It is better to ask a lot of questions with limited answer options than to ask a small number of questions where students can give long answers. Try and ensure a balance between the length of the answers and the time available for the students.
- Don't use words that express a possibility, such as 'might', 'can', 'may', 'could be', etc.
- You should only use definite cardinal numbers. In other words, don't use phrases like: 'State *a number of* characteristics'. You should phrase this as follows: 'State *three* characteristics ...'.

- Avoid phrasing questions so that they can be answered with Yes or No. In other words, don't use phrases like: 'Can you name three reasons why ...'. You should phrase this as follows: 'Name three reasons why ...'
- Avoid phrasing questions in the negative. In other words, don't use phrases like: 'Which persons do not agree with the following propositions ...?'. You should phrase this as follows: 'Name the persons opposed to the propositions alleging that ...'.
- Make it clear whether the question is asking for the student's own opinion or the opinion of someone else. In other words, don't use phrases like: 'Name four points of criticism on ...'. You should phrase this as follows: 'What four points of criticism does the author of the book make about ...'.
- Be careful when using phrases taken from the subject matter. Always check the meaning of the question and then rephrase it in your own words.
- Never ask two questions in the same sentence.
- Avoid the use of difficult words, except those that the students have to know in order to show that they have mastered the subject matter.
- Always observe all rules relating to grammar, spelling and style.
- Phrase your questions in such a specific way that all the experts are able to agree on which answer is better than the others. Questions should not be phrased in such general terms that students can answer them in any way they please.
- State the assignments (what you are asking students to do) as comprehensively and specifically as possible. It must be absolutely clear to the students how much detail is required and how long their answers may be.
- Make sure the students have sufficient time to answer each of the questions.
- Phrase your questions in such a way that no misunderstandings can arise. Questions must be specific.
- Phrase your questions in the form of statements as far as possible. This will result in more direct answers than if you phrase them in the form of questions. Example: 'Mintzberg distinguishes three management levels at organisations: strategic level, organisational level and operational level. Discuss the relationship between these levels.'

Appendix 4: Standardisation schedules for multiple-choice questions

Standardisation of 40 four-choice questions			
Probability of guesswork = 10; Awarding of points: 30 x 0.33333			
State the following when processing an examination: P1: 10 = 0 and P2: 40 = 10.			
Enter the 'total number of points' with two figures after the decimal point in order to avoid problems with rounding off.			
Use 'total number of points' for processing as a partial grade.			
Use 'mark' if the results are entirely determined by the multiple-choice questions.			
number of correct answers	total number of points		mark
1	0.00		1
2	0.00		1
3	0.00		1
4	0.00		1
5	0.00		1
6	0.00		1
7	0.00		1
8	0.00		1
9	0.00		1
10	0.00		1
11	0.33		1
12	0.67		1
13	1.00		1
14	1.33		1
15	1.67		2
16	2.00		2
17	2.33		2
18	2.67		3
19	3.00		3
20	3.33		3
21	3.67		4
22	4.00		4
23	4.33		4
24	4.67		5
25	5.00		5
26	5.33		5

27	5.67		6
28	6.00		6
29	6.33		6
30	6.67		7
31	7.00		7
32	7.33		7
33	7.67		8
34	8.00		8
35	8.33		8
36	8.67		9
37	9.00		9
38	9.33		9
39	9.67		10
40	10.00		10

If there are 60 three-choice questions, the probability of guesswork is 20. In that case, you should state the following when processing the examination: P1: 20 = 0; P2: 60 = 10. The value of the 21st and all the following correct answers will then be 0.25 points: $40 \times 0.25 = 10$. In this case, there will be no rounding-off problems, which means that the ‘total number of points’ can be stated exactly to 1 figure after the decimal point.

If there are 80 two-choice questions, the probability of guesswork is 40. In that case, you should state the following when processing the examination: P1: 40 = 0; P2: 80 = 10. The value of the 21st and all the following correct answers will be 0.25 points here as well: $40 \times 0.25 = 10$.

When deleting one or more questions, make sure you leave them out when processing the marks. Establish the appropriate probability of guesswork for all the remaining questions, and adjust the values for P1 and P2 accordingly. For example, if one question is deleted:

- If there are 40 four-choice questions, there is no need to adjust the probability of guesswork: one-quarter of 39 = 10. Deduct 1 point from P1 and P2.
- If there are 60 three-choice questions, there is no need to adjust the probability of guesswork: one-third of 59 = 20. Deduct 1 point from P1 and P2.
- If there are 80 two-choice questions, you should adjust the probability of guesswork to 39: one-half of 79 = 39. Deduct 1 point from P1 and P2.

After adjusting the number of points awarded, check what happens to the automatic rounding off and allocation of marks. Adjust this if necessary.