English

Dansk
Deutsch
Ελληνικα
English
Español
Suomi
Français
Gaeilge
Islenska
Italiano
Nederlands
Norsk
Português
Svenska

DIALANG

INDEX

DIALANG

Introduction to DIALANG

Context

DIALANG is a computer-based and computer-scored suite of diagnostic tests of foreign language abilities in 14 European languages. There are tests of Reading, Writing, Listening, Structures and Vocabulary, in Danish, Dutch, English, Finnish, French, German, Greek, Icelandic, Irish, Italian, Norwegian, Portuguese, Spanish and Swedish. In addition, there is a Vocabulary Size Placement Test for each language, and, also for each language, there are self-assessment 'I Can' statements for Reading, Writing and Listening. DIALANG provides extensive feedback to users, and advice on how they might improve their abilities.

The DIALANG Project was funded by the European Commission under the Socrates LINGUA programme, ACTION D. The initial development of DIALANG, beginning at the end of 1996, was the result of collaboration among over 20 partner institutions, most of them universities, throughout Europe, coordinated by the University of Jyväskylä. For full details, visit the DIALANG website at www.dialang.org. The piloting, calibration and further development of the system was coordinated by the Free University of Berlin. All tests can be downloaded from the DIALANG website, free of charge, and can be taken over the Internet, again free of charge, from any computer on which the program is installed.

DIALANG is intended to help learners diagnose their strengths and weaknesses in any or all of the five abilities tested, to encourage learners to self-assess their abilities, and to compare their self-assessment with their test results. By so doing it is intended to enhance learner autonomy by encouraging learner reflection on what they know and can do in their target foreign language, and thereby to increase their awareness of what is involved in learning or improving in a foreign language.

The DIALANG Project began in 1996, when the first draft of the Council of Europe's Common European Framework of Reference (CEFR) was already available, and so it was possible to base the DIALANG Assessment Framework and Specifications on the CEFR. DIALANG also used the 'Can Do' statements contained in the CEFR, but adjusted their wording to suit its self-assessment purposes, essentially by changing 'Can-Do' into 'I Can' and at times simplifying the language to make it more accessible to learners. In addition, extensive explanations of the feedback given by the DIALANG system were based on the CEFR scales, and the advice provided to learners on how to improve their ability to use the language used the information contained in the CEFR's Descriptive Scheme (which became Chapters 4, 5, 6, and 7 in the 2001 publication - Council of Europe, 2001).

DIALANG not only contains tests in its 14 languages, but it also enables users to read test instructions and rubrics, help facilities, introductions and, most importantly, all self-assessment statements, feedback and advice, in any of these 14 languages. Once agreement had been reached on the appropriate wording, particularly of the 'I Can' statements, and the Explanatory and Advisory Feedback, these were then translated from the original English into the 13 languages of DIALANG, and checked for the quality of the translations. Thus DIALANG is multilingual in both the tests available, and in the interface and administration of the system.

DIALANG

Although intended to offer detailed diagnosis of language skills and subskills (for full details of these, see Alderson, 2005), DIALANG also reports the learner's ability in the skill or aspect of language being tested, as a level on the CEFR (A1 to C2), rather than as a numerical score.

DIALANG is intended to contribute to life-long learning and learner autonomy, and thus is "aimed" at any adult learner of foreign languages, whether or not that learner is engaged in institution-based language learning. However, many institutions consider the DIALANG system to be of value in placing students into language classes which are based on the CEFR, and thus, although intended to be diagnostic, the tests are increasingly being used for placement purposes. At the time of writing, roughly 1,500 users take a DIALANG test every day.

Test development process

a) Content quality

The DIALANG tests were developed by Assessment Development Teams (ADTs) for each of the 14 languages. Typically a team was composed of language educationalists, usually at institutions of adult or higher education. Members were experienced teachers, researchers and assessment experts, where these existed for the language in question. In several cases, where the traditions of assessment were different from those espoused in DIALANG and the CEFR, familiarisation with the CEFR and with appropriate methods of test and item writing was made available and a number of training sessions were held. The tests themselves, as already noted, were based on the DIALANG Assessment Framework (DAF) and DIALANG Assessment Specifications (DAS), themselves based on or derived from the CEFR and related User Guides. Detailed Guidelines for Item Writers were also developed and made available as the formats (test methods) of the tests and an associated computer-based authoring system became available.

Test development was coordinated by two Test Development Coordinators, who were responsible for the issue and updating of all documentation and guidance, and who communicated with the ADTs through the ADT Team Leader. As items were developed, first on paper and then over the Internet using the Authoring software developed by the Project, they were subject to an extensive process of quality review, first by peers within the ADT, then by the Team Leader, and then by external reviewers commissioned by the Test Development Coordinators, using the Internet-based Item Review software. Once revisions had been made and reviewed by the ADT, the Test Development Coordinators made a second review to ensure that all items conformed to the DAF and DAS, and to the requirements of the computer programs. In this way, large pools of items were developed for each skill for each language. Finally, in preparation for empirical piloting, ADT Team leaders were requested to select the best 300 items from their own item pool for the first round of piloting.

b) Piloting and calibration

In the second phase of the Project, arrangements were made for the computer-based piloting of items using the specially developed Pilot Tool software also developed by the Project (in the first phase of the Project, Finnish had been piloted in paper and pencil form, in order to try out both the overall pilot design and the calibration analyses). Initially a piloting design was implemented where 12 booklets would be piloted, containing 30 items from a skill (Reading, Writing, Listening) plus 20 items from a language aspect (Structures, Vocabulary), with each item appearing in two separate booklets, as well as 150 items from the Vocabulary Size Placement Test (VSPT) and between 30 and 40 self-assessment (SA) statements. Booklets were assigned randomly to pilot test-takers. This design would enable IRT-based item calibrations

DIALANG

to take place, using the computer program OPLM, once 600 learners had taken a pilot test (implying 100 responses per item) and a second calibration was to take place once 1200 users had responded (200 responses per item).

However, once sufficient numbers of responses to the VSPT and SAs had been gathered to enable calibrations, the pilot booklet design was modified to reduce the number of VSPT items being tested to 99 and the SA statements to 18 per skill. This allowed more test items to be included in each booklet, and only 450 test-takers would be required before an initial item calibration was possible. To date, initial calibrations have been possible for Finnish, English, French, German and Spanish, and Italian, Swedish and Dutch are getting close to the target number. English has undergone a second calibration, whose results were remarkably similar to the first round, suggesting a second calibration may not be necessary for other languages.

Standard Setting

Two different methods of standard setting were developed in DIALANG, and the Project developed guidelines which were subsequently published in the Manual and the Supplement. Details are available in Alderson (2005). In essence, the first method was a modified Angoff procedure, where probablity judgements by up to 12 experts were replaced by Yes-No decisions (numbers of judges varied by language and skill).

"The task for the judge is to state for each item in the test

At what CEF level can a test taker already answer the following item correctly?"
...
In practice the procedure follows a pattern also presented in the Manual:
- Familiarisation with the CEF
- Training (with feedback) on how to judge ... new items ... in terms of CEF levels;
- Judgement of items to be related to the CEF

(Manual, page 91)

This method was applied to Finnish, English and Spanish, and detailed results are reported in Alderson (2005).

Partly because of the lack of an adequate method of judging the reliability and consistency of judges, a second process was developed which involved experts in making two sets of judgements. The first round required them to place items into piles, representing the different levels of the CEFR, starting with A1, in answer to the question:

"Is it reasonable to require that a learner at level A1 gets this item right?"

If yes, then the item belongs to A1. If not, then it is a candidate for the next level up. Once piles have been created, they are checked for internal consistency and items may be moved if necessary.

After a period of time, minimally two to three hours, the process is repeated, thereby enabling intra-rater consistency to be calculated alongside the inter-rater agreement. Detailed results are available in Alderson (2005). Only reliable judges are retained in the subsequent analysis.

Finally, once item calibration details are available, the empirical results are combined with the judged CEFR level for each item, in a computer program specially developed by the Project.

References

Alderson, J. C. (2005) *Diagnosing foreign language proficiency: the interface between learning and assessment*. London: Continuum Books

Council of Europe (2003) *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF). Manual: Preliminary Pilot Version*. Strasbourg: Language Policy Division, Council of Europe.

![DIALANG logo]

Information about the Dialang Items

The items presented in this section are taken from Version 1 of the freely available English test and represent a selection across the different CEFR levels. Each item is accompanied by a summary of its content taken from the Dutch CEF Grid (www.ling.lancs.ac.uk/cefgrid), itself closely based on the CEFR. In addition each item is accompanied by information on its statistical properties.

Statistical information

The calibration and standard setting processes described in the introduction to Dialang resulted in information about the properties of each item in terms of classical test theory - in particular, item facilities and discriminations. In addition, IRT analyses resulted in estimations of the ability of learners getting a given score, in terms of theta estimates of ability. Cut-offs on this theta scale are based on the standard-setting judgements, and an item's logit value can be related to the logit values for ability levels.

Guideline To Information On DIALANG items

Below follows an explanation of the statistical and other information accompanying the se-lected DIALANG items:

— the Framework level of the item is indicated immediately before the screenshot of the item
— the subskill of listening that the item was designed to measure is indicated immediately after the screenshot; DIALANG listening items aim at tapping one of these three skills:
  o Identifying the main idea(s) or information in the spoken discourse.
  o Listening intensively for specific detail.
  o Inferencing; i.e., the ability to infer meaning on the basis of spoken discourse.

The table under 'Information from piloting' describes the results for the booklets where the item was piloted.
— booklet number = the number of the pilot booklet in which the item was piloted (each DIALANG item appeared in two different booklets in the pilot design)
— n size = the number of pilot test takers for the booklet in question
— Facility = percent of pilot test takers who replied the item correctly
— Rpbis (unweighted) = point-biserial correlation between the item and the other listening items in the pilot booklet when all items have equal weight
— Rpbis (weighted) = point-biserial correlation between the item and the other listening items in the pilot booklet when the items are weighted by their discrimination index ('A' in the table under 'Results of calibration')

The table 'Results of calibration' display the key statistical information for the item from the OPLM analysis.

nr = number of the item in the OPLM analysis

label = label of the item (which in fact is also the item ID in the DIALANG system)

A = discrimination index

B = theta value of the item (which indicates its difficulty / position on the scale after the items have been scaled; the cut-offs for the Framework levels are expressed on this same scale)

SE(B) = standard error of the theta value

S = S test; in this the respondents are grouped into two or more (max 8) groups of equal size and the ability of the item to discriminate the groups is analysed

DF = degrees of freedom in the S analysis

P = probability of the S test (this in fact indicates whether the item studied fits the overall model)

M, M2, M3 = other indicators of the model fit for the item

> — M = in this analysis, two groups are formed of the test takers and it is analysed whether the item can discriminate the two groups (the low group are those whose probability to get the item right is at most .4, whereas in the high group the probability is .6 or higher)

> — M2 = in this analysis, respondents are split into two halves according to their overall score in the test / booklet

> — M3 = in this analysis, the respondents are grouped into three groups of equal size, according to their test score

P or M-value of 99.999 indicates that a proper value cannot be computed because the respondents cannot be divided into equal groups; this occurs with very easy or very difficult items. For the same reason, the S test value can sometimes be .000.

An M value of the size of + or – 2 or above indicates that there may be a problem in the item (it cannot properly discriminate test takers at different ability levels)

Analysis of item content

The items on this CD have been analysed using the Dutch CEF Grid produced by the Dutch CEF Construct Project. This Project aimed to help test developers and other language educationists construct or relate test items to the CEFR. (The Final Report of the Project has been included on this CD for ease of reference.) One major outcome of the Project was an Internet-based Grid which can be used to help characterise reading and listening texts, items and tasks. Application of the Grid results in summary tables and the Grid has been used with the DIALANG items to produce the summary tables below. (Readers of this Users' Guide can access the Grid at:

www.ling.lancs.ac.uk/cefgrid

and can use it in the content analysis of their own texts, items and tasks.)

DIALANG

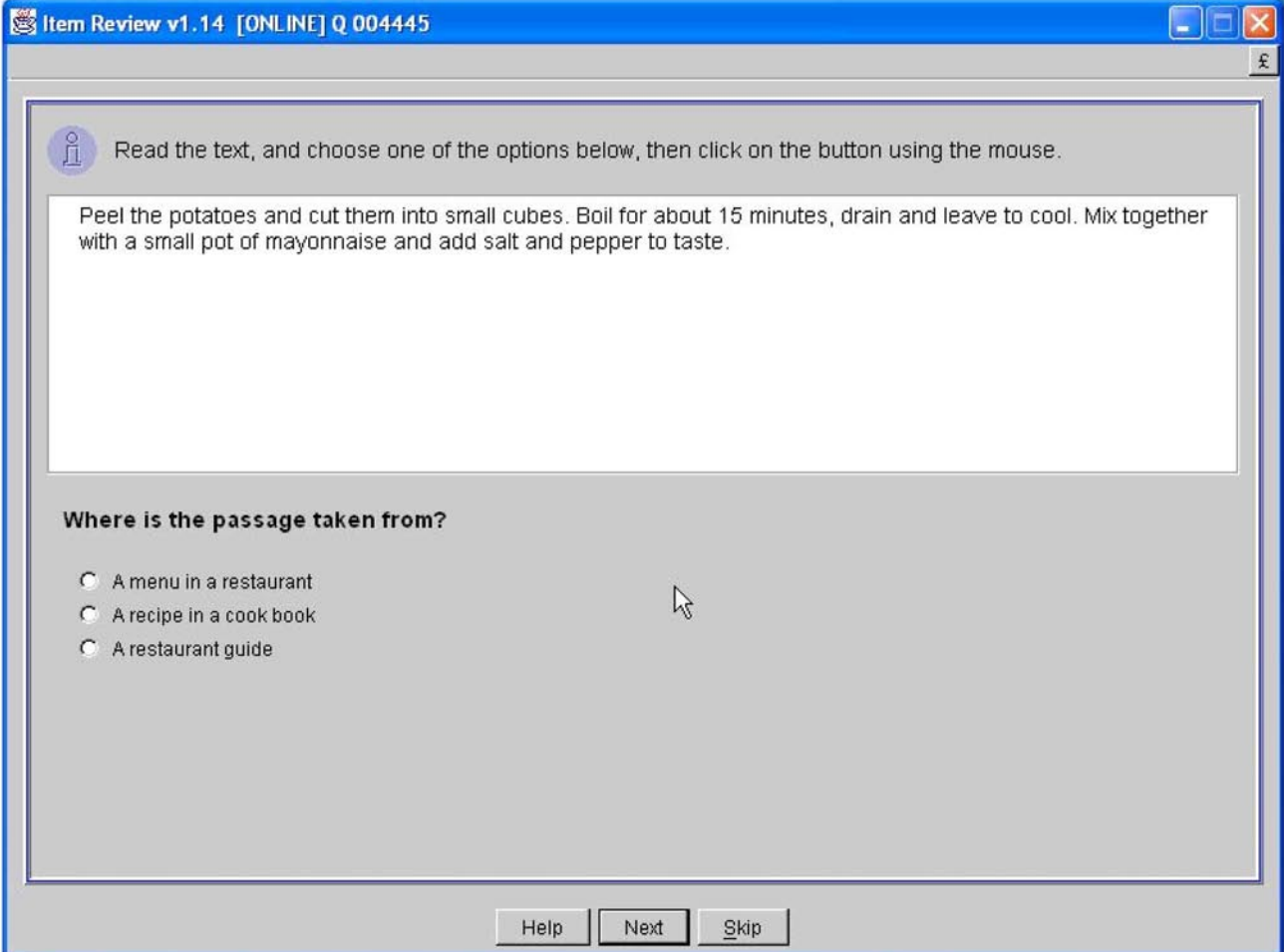# Dialang English Reading Comprehension Items

General information

There were 349 – 370 pilot test takers per each reading item. Overall data – model fit was satisfactory (p = .83).

The cut-offs on the theta scale for the CoE levels are based on standard setting procedures where 9 – 11 experts judged the items (# of experts varied from skill to skill); the standard setting data were analysed with a special programme designed by Norman Verhelst at CITO.

Cut-off points for the Framework levels for the DIALANG English reading items:

under A1: theta under -.94
A1: theta between -.94 and -.56
A2: theta between -.56 and -.22
B1: theta between -.22 and .21
B2: theta between .21 and 1.23
C1: theta between 1.23 and 2.01
C2: theta above 2.01

DIALANG

A1

Item Review v1.14 [ONLINE] Q 004445

ℹ️   Read the text, and choose one of the options below, then click on the button using the mouse.

Peel the potatoes and cut them into small cubes. Boil for about 15 minutes, drain and leave to cool. Mix together with a small pot of mayonnaise and add salt and pepper to taste.

Where is the passage taken from?

○ A menu in a restaurant
○ A recipe in a cook book
○ A restaurant guide

Help   Next   Skip

Content analysis

Text Characteristics:

Task:	4445
1. Text Source:	Recipes
2. Authenticity:	Authentic
3. Discourse type:	Mainly Instructive
4. Domain:	Personal
5. Topic:	10. Food and drink
6. Nature of Content:	only concrete content
7. Text Length:	34
8. Vocabulary:	mostly frequent vocabulary
9. Grammar:	only simple structures
Comprehensible by learner at CEF level:	A2

Item Characteristics:

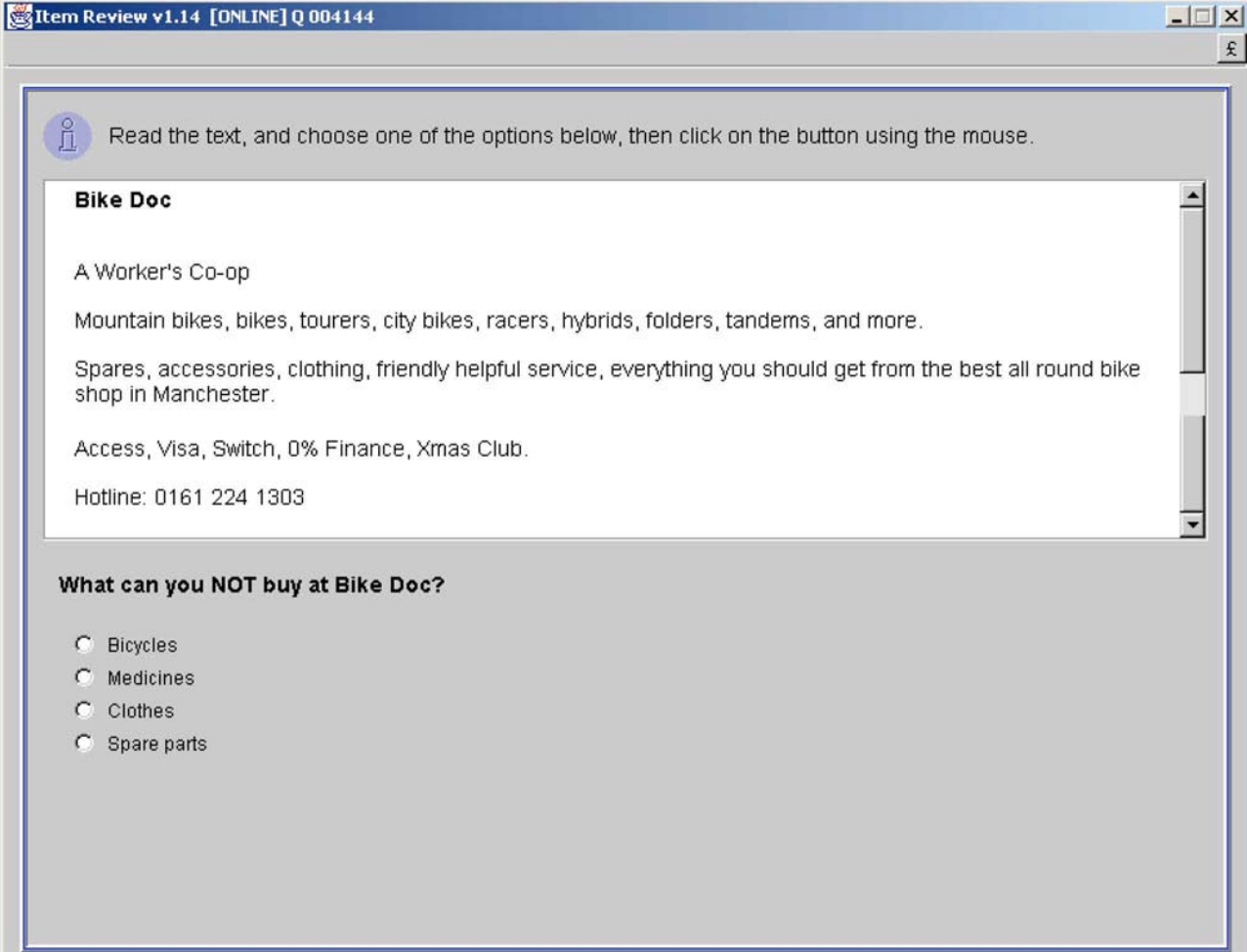| Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Recognise and Retrieve
15. Operations:	Explicit
15. Operations:	Communicative Purpose
Item Level Estimated:	A1

Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
1	183	.989	.227	.248
4	187	.989	.255	.275

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
70 R004445	5	-.834	.102	.000	0	99.999	99.999	-.298	-.564

A1 (close to A2)

Item Review v1.14 [ONLINE] Q 004144

ℹ️ Read the text, and choose one of the options below, then click on the button using the mouse.

Bike Doc

A Worker's Co-op

Mountain bikes, bikes, tourers, city bikes, racers, hybrids, folders, tandems, and more.

Spares, accessories, clothing, friendly helpful service, everything you should get from the best all round bike shop in Manchester.

Access, Visa, Switch, 0% Finance, Xmas Club.

Hotline: 0161 224 1303

What can you NOT buy at Bike Doc?

- ○ Bicycles
- ○ Medicines
- ○ Clothes
- ○ Spare parts

Content analysis

Text Characteristics:

Task:	4144
Skill:	reading
1. Text Source:	Advertising material
2. Authenticity:	Authentic
3. Discourse type:	Mainly Descriptive
4. Domain:	Public
5. Topic:	9. Shopping
6. Nature of Content:	mostly concrete content
7. Text Length:	48
8. Vocabulary:	rather extended
9. Grammar:	only simple structures
Comprehensible by learner at CEF level:	A2/B1

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Make inferences
15. Operations:	Explicit
15. Operations:	Detail
Item Level Estimated:	A1/A2

Statistical analysis
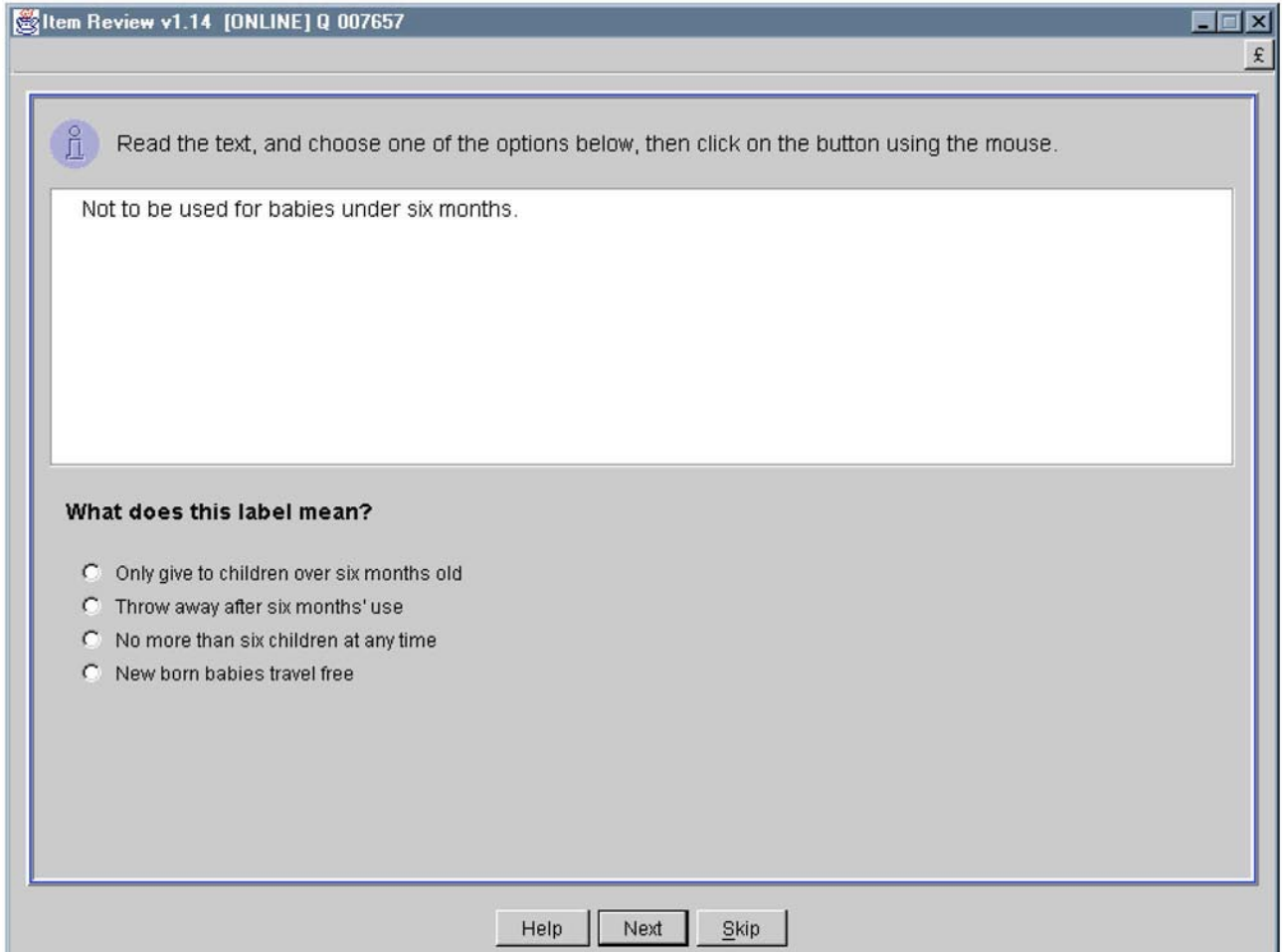
Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
1	183	.940	.391	.438
2	171	.947	.362	.397

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
83 R004144	4	-.565	.063	3.961	1	.047	-.915	.663	.220

A2

Item Review v1.14 [ONLINE] Q 007657

Read the text, and choose one of the options below, then click on the button using the mouse.

Not to be used for babies under six months.

What does this label mean?

○ Only give to children over six months old
○ Throw away after six months' use
○ No more than six children at any time
○ New born babies travel free

Help   Next   Skip

Content analysis

Text Characteristics:

Task:	7657
1. Text Source:	Labeling and packaging
2. Authenticity:	Authentic
3. Discourse type:	Mainly Instructive
4. Domain:	Public
5. Topic:	7. Health and bodycare
6. Nature of Content:	only concrete content
7. Text Length:	9
8. Vocabulary:	only frequent vocabulary
9. Grammar:	Limited range of complex structures
Comprehensible by learner at CEF level:	A2

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Recognise and Retrieve
15. Operations:	Explicit
15. Operations:	Main idea/gist
Item Level Estimated:	A2

Statistical analysis
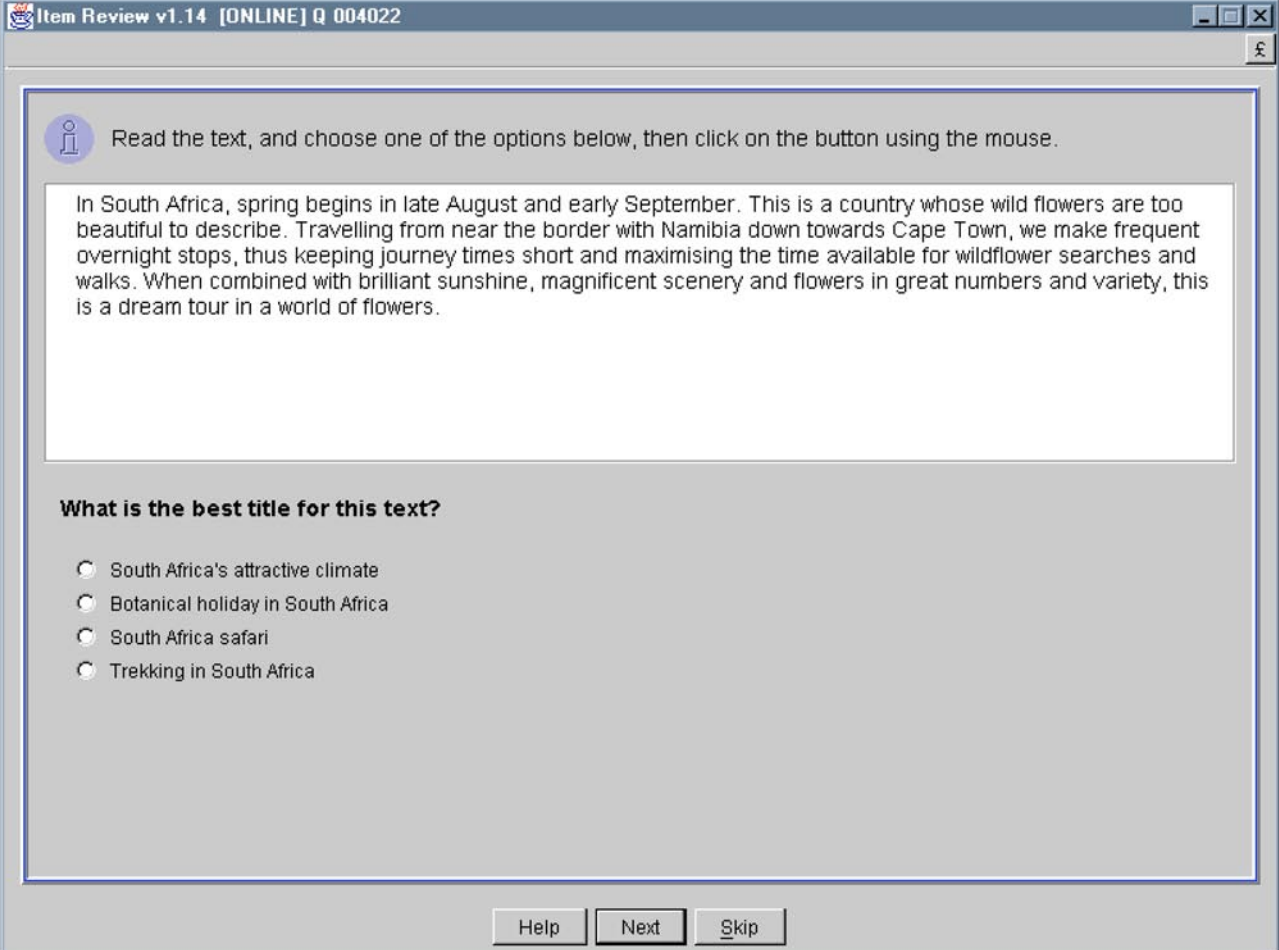
Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
3	178	.910	.537	.578
4	187	.930	.456	.524

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
107 R007657	6	-.337	.041	.339	1	.560	-.315	-.486	-1.054

A2

Item Review v1.14 [ONLINE] Q 004022

Read the text, and choose one of the options below, then click on the button using the mouse.

In South Africa, spring begins in late August and early September. This is a country whose wild flowers are too beautiful to describe. Travelling from near the border with Namibia down towards Cape Town, we make frequent overnight stops, thus keeping journey times short and maximising the time available for wildflower searches and walks. When combined with brilliant sunshine, magnificent scenery and flowers in great numbers and variety, this is a dream tour in a world of flowers.

What is the best title for this text?

- South Africa's attractive climate
- Botanical holiday in South Africa
- South Africa safari
- Trekking in South Africa

Help   Next   Skip

Content analysis

Text Characteristics:

Task:	4022
1. Text Source:	Brochures
2. Authenticity:	Authentic
3. Discourse type:	Mainly Descriptive
4. Domain:	Public
5. Topic:	5.Travel
6. Nature of Content:	mostly concrete content
7. Text Length:	78
8. Vocabulary:	mostly frequent vocabulary
9. Grammar:	Limited range of complex structures
Comprehensible by learner at CEF level:	B1

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Evaluate
15. Operations:	Explicit
15. Operations:	Main idea/gist
Item Level Estimated:	B1

Statistical analysis
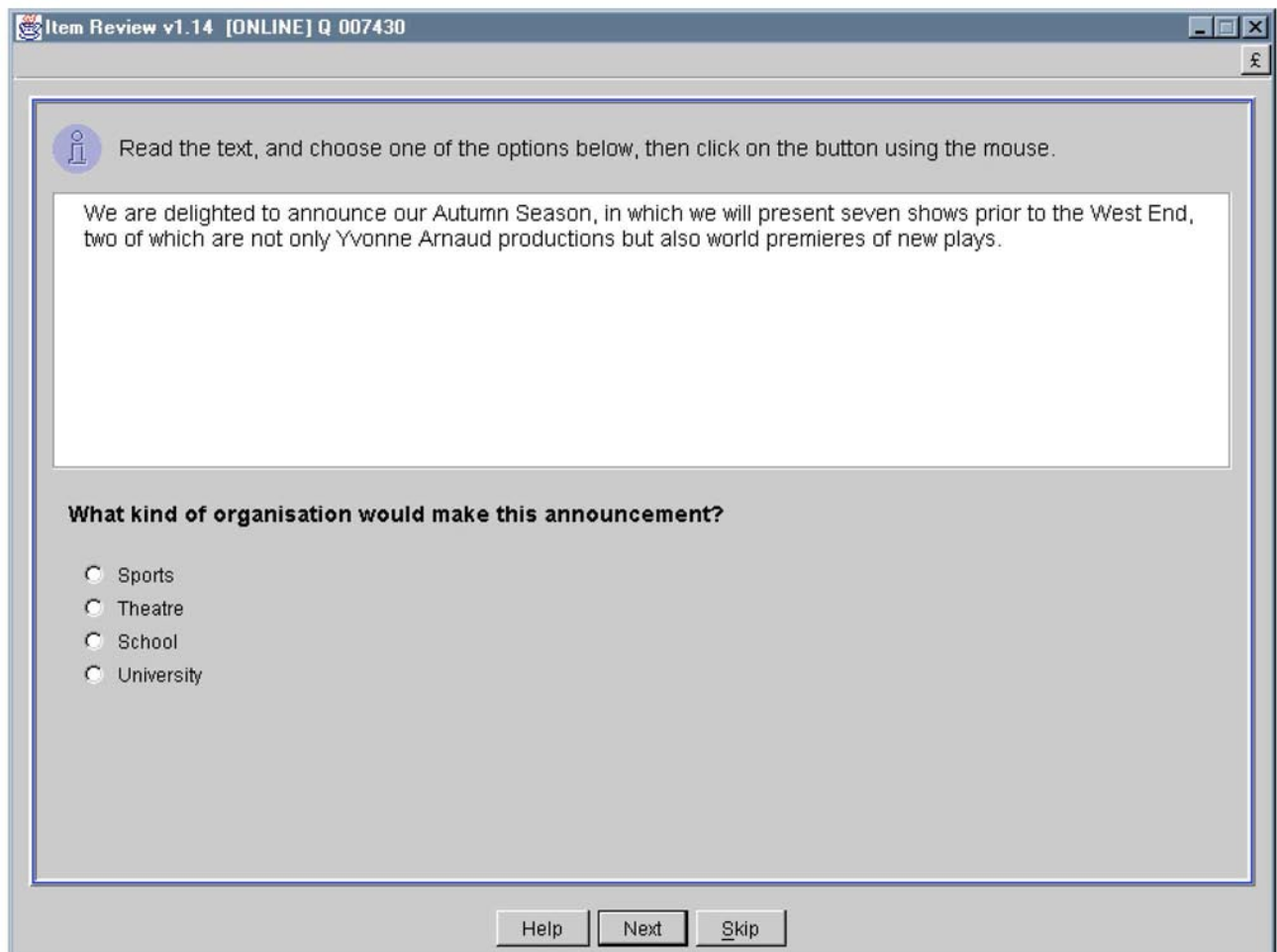
Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
1	183	.781	.409	.362
4	187	.765	.436	.387

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
63 R004022	2	-.308	.065	3.675	6	.721	.067	-.985	-.830

A2

Item Review v1.14 [ONLINE] Q 007430

Read the text, and choose one of the options below, then click on the button using the mouse.

We are delighted to announce our Autumn Season, in which we will present seven shows prior to the West End, two of which are not only Yvonne Arnaud productions but also world premieres of new plays.

What kind of organisation would make this announcement?

- ○ Sports
- ○ Theatre
- ○ School
- ○ University

Help   Next   Skip

Content analysis

Text Characteristics:

Task:	7430
1. Text Source:	Advertising material
2. Authenticity:	Authentic
3. Discourse type:	Mainly Instructive
4. Domain:	Public
5. Topic:	4. Free time, entertainment
6. Nature of Content:	mostly concrete content
7. Text Length:	36
8. Vocabulary:	mostly frequent vocabulary
9. Grammar:	Limited range of complex structures
Comprehensible by learner at CEF level:	B1

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Recognise and Retrieve
15. Operations:	Explicit
15. Operations:	Detail
Item Level Estimated:	A2/B1

Statistical analysis
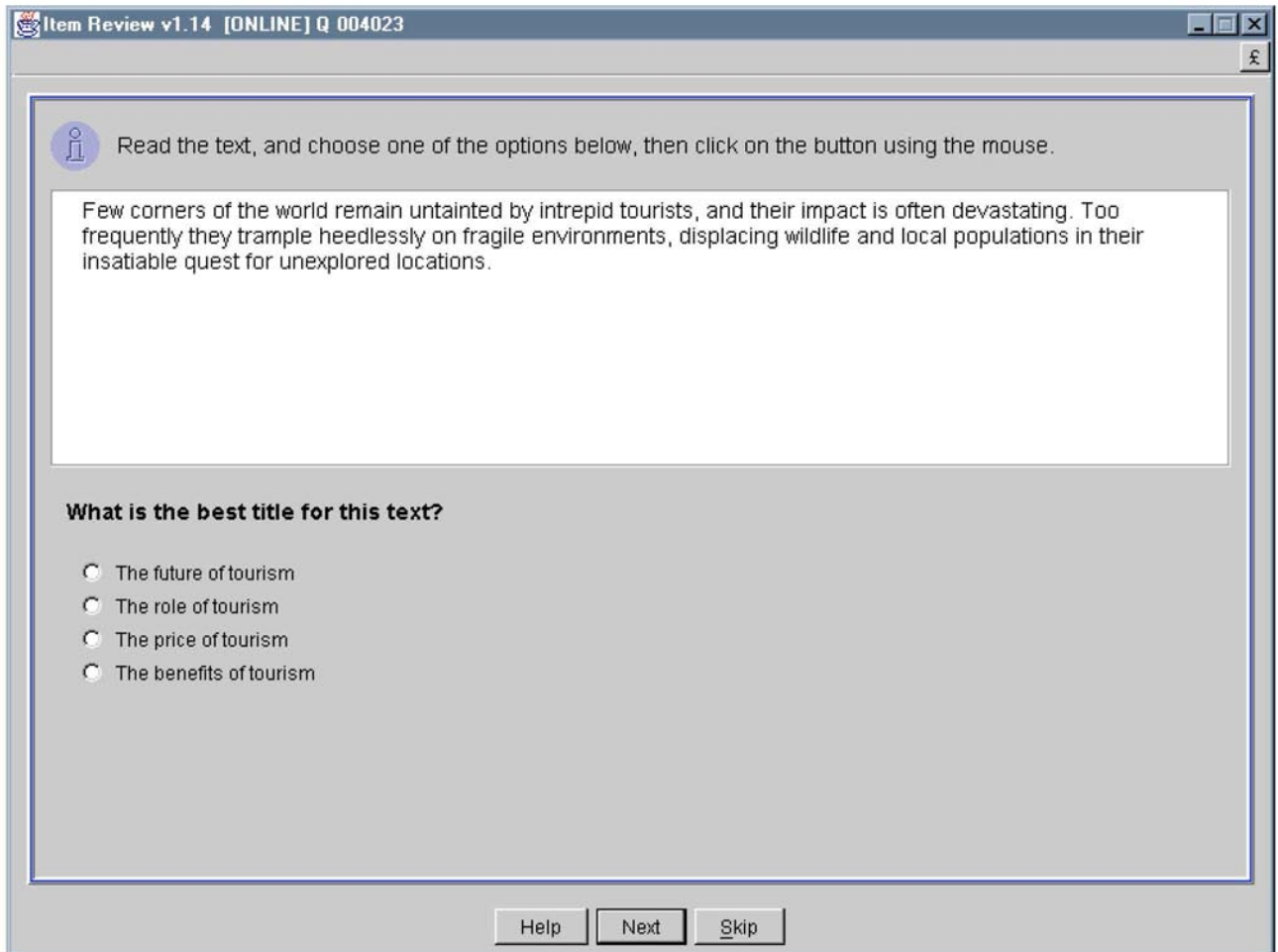
Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
1	183	.956	.432	.497
2	171	.942	.548	.590

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
86 R007430	6	-.439	.049	.000	0	99.999	-.466	-.499	-1.014

DIALANG

B1

Item Review v1.14 [ONLINE] Q 004023

Read the text, and choose one of the options below, then click on the button using the mouse.

Few corners of the world remain untainted by intrepid tourists, and their impact is often devastating. Too frequently they trample heedlessly on fragile environments, displacing wildlife and local populations in their insatiable quest for unexplored locations.

What is the best title for this text?

- ○ The future of tourism
- ○ The role of tourism
- ○ The price of tourism
- ○ The benefits of tourism

Help   Next   Skip

Content analysis

Text Characteristics:

Task:	4023
1. Text Source:	Magazines
2. Authenticity:	Authentic
3. Discourse type:	Mainly Argumentative
4. Domain:	Public
5. Topic:	5.Travel
6. Nature of Content:	mainly abstract content
7. Text Length:	36
8. Vocabulary:	rather extended
9. Grammar:	Limited range of complex structures
Comprehensible by learner at CEF level:	B2

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Evaluate
15. Operations:	Implicit
15. Operations:	Main idea/gist
Item Level Estimated:	B2

Statistical analysis
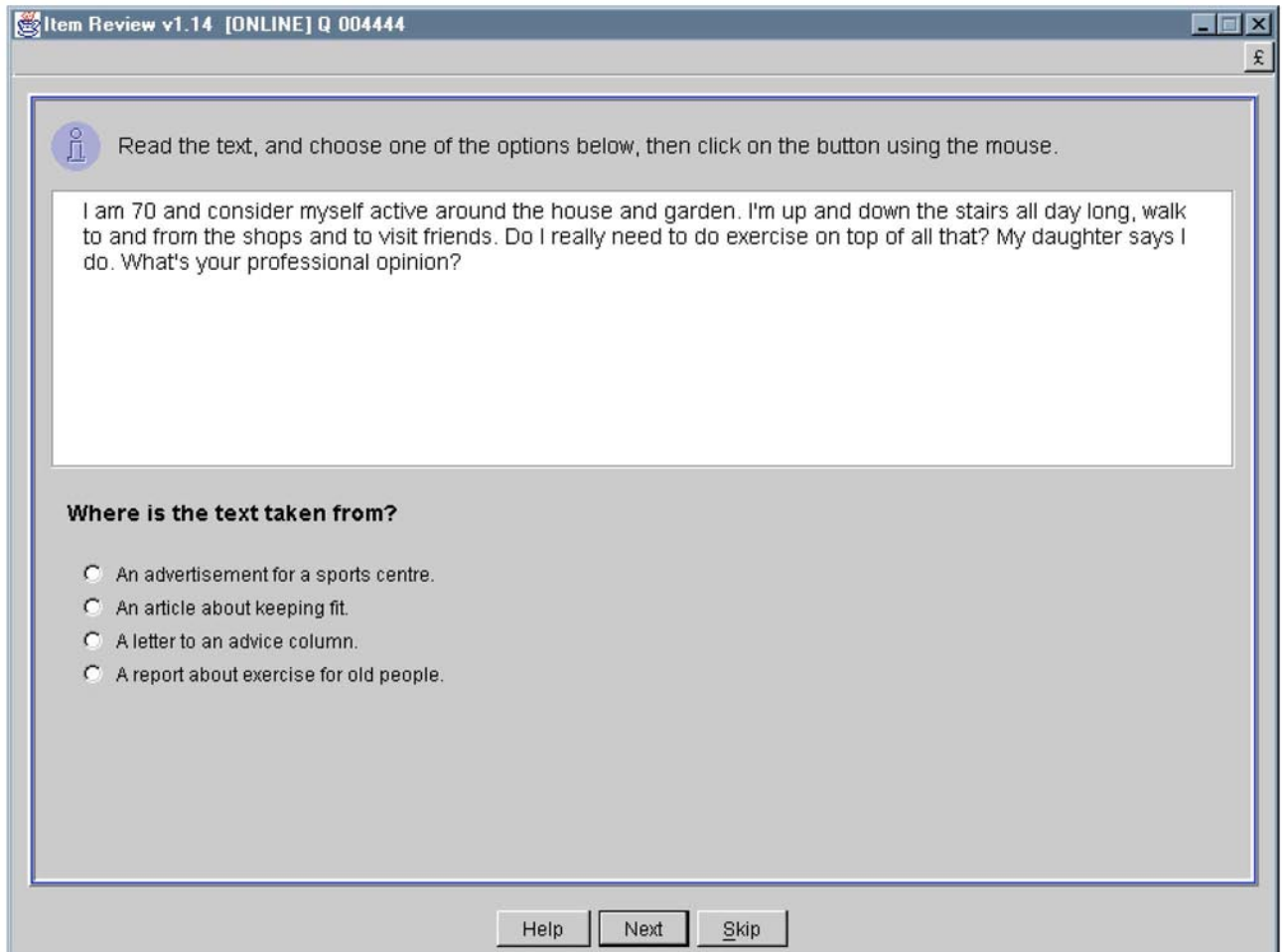
Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
1	183	.787	.663	.708
2	171	.801	.550	.587

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
76 R004023	5	-.057	.035	1.263	4	.868	-.729	.183	-.428

DIALANG

B1

Item Review v1.14 [ONLINE] Q 004444

ℹ️ Read the text, and choose one of the options below, then click on the button using the mouse.

I am 70 and consider myself active around the house and garden. I'm up and down the stairs all day long, walk to and from the shops and to visit friends. Do I really need to do exercise on top of all that? My daughter says I do. What's your professional opinion?

Where is the text taken from?

○ An advertisement for a sports centre.

○ An article about keeping fit.

○ A letter to an advice column.

○ A report about exercise for old people.

| Help | Next | Skip |

Content analysis

Text Characteristics:

Task:	4444
1. Text Source:	Magazines
2. Authenticity:	Authentic
3. Discourse type:	Mainly Descriptive
4. Domain:	Personal
4. Domain:	Public
5. Topic:	7. Health and bodycare
6. Nature of Content:	mostly concrete content
7. Text Length:	54
8. Vocabulary:	mostly frequent vocabulary
9. Grammar:	only simple structures
Comprehensible by learner at CEF level:	A2/B1

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Make inferences
15. Operations:	Explicit
15. Operations:	Communicative Purpose
Item Level Estimated:	A2/B1

Statistical analysis
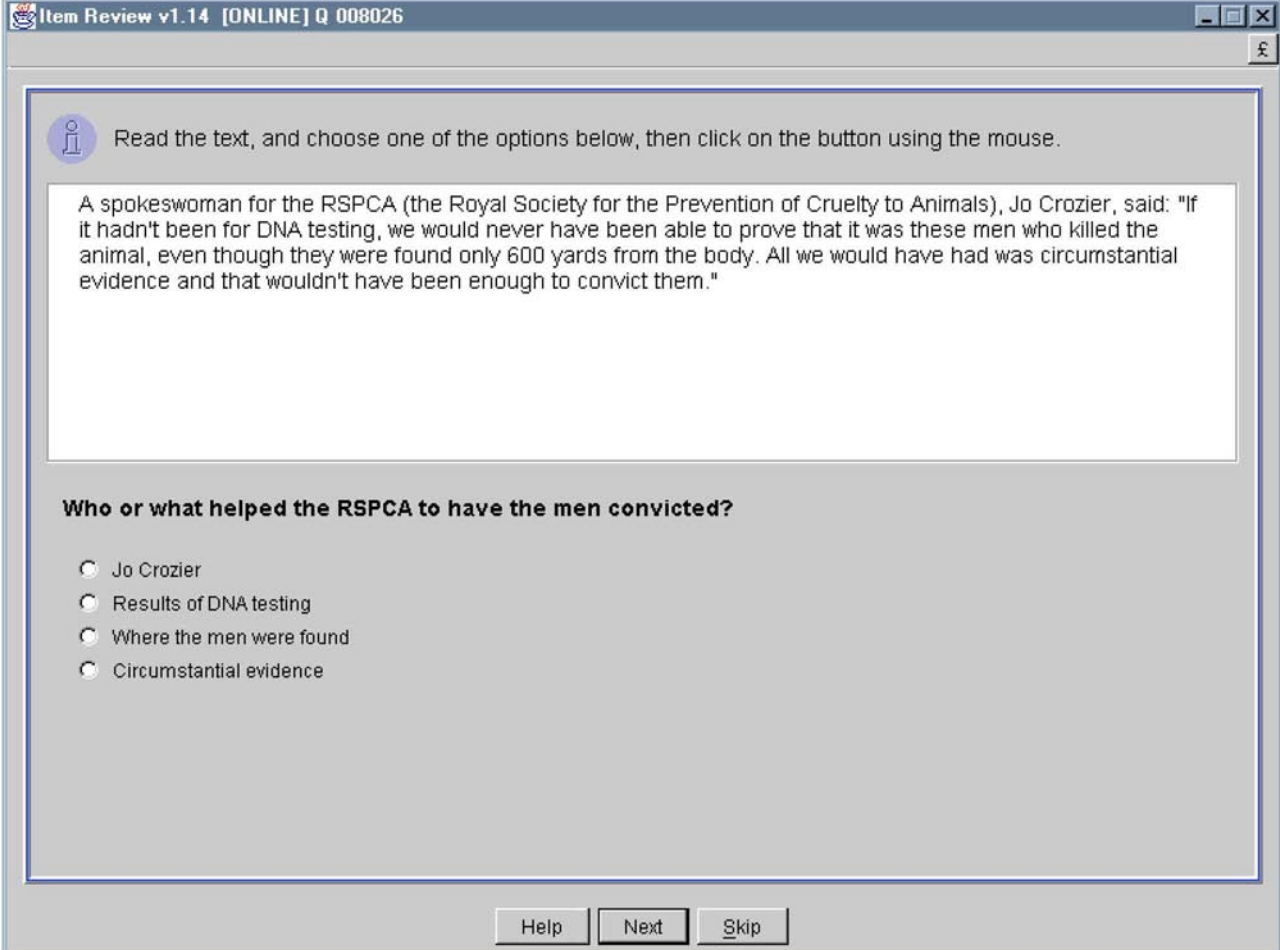
Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
2	171	.743	.585	.601
3	178	.730	.595	.615

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
101 R004444	4	.026	.039	14.635	5	.012	-.105	-1.061	-2.270

B1

Item Review v1.14 [ONLINE] Q 008026

Read the text, and choose one of the options below, then click on the button using the mouse.

A spokeswoman for the RSPCA (the Royal Society for the Prevention of Cruelty to Animals), Jo Crozier, said: "If it hadn't been for DNA testing, we would never have been able to prove that it was these men who killed the animal, even though they were found only 600 yards from the body. All we would have had was circumstantial evidence and that wouldn't have been enough to convict them."

Who or what helped the RSPCA to have the men convicted?

○ Jo Crozier
○ Results of DNA testing
○ Where the men were found
○ Circumstantial evidence

Help   Next   Skip

Content analysis

Text Characteristics:

Task:	8026
1. Text Source:	Newspapers
2. Authenticity:	Authentic
3. Discourse type:	Mainly Narrative
4. Domain:	Public
5. Topic:	15. Other
Specify other topic:	Crime
6. Nature of Content:	mostly concrete content
7. Text Length:	70
8. Vocabulary:	rather extended
9. Grammar:	Wide range of complex structures
Comprehensible by learner at CEF level:	B2

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Recognise and Retrieve
15. Operations:	Explicit
15. Operations:	Detail
Item Level Estimated:	B1/B2

Statistical analysis
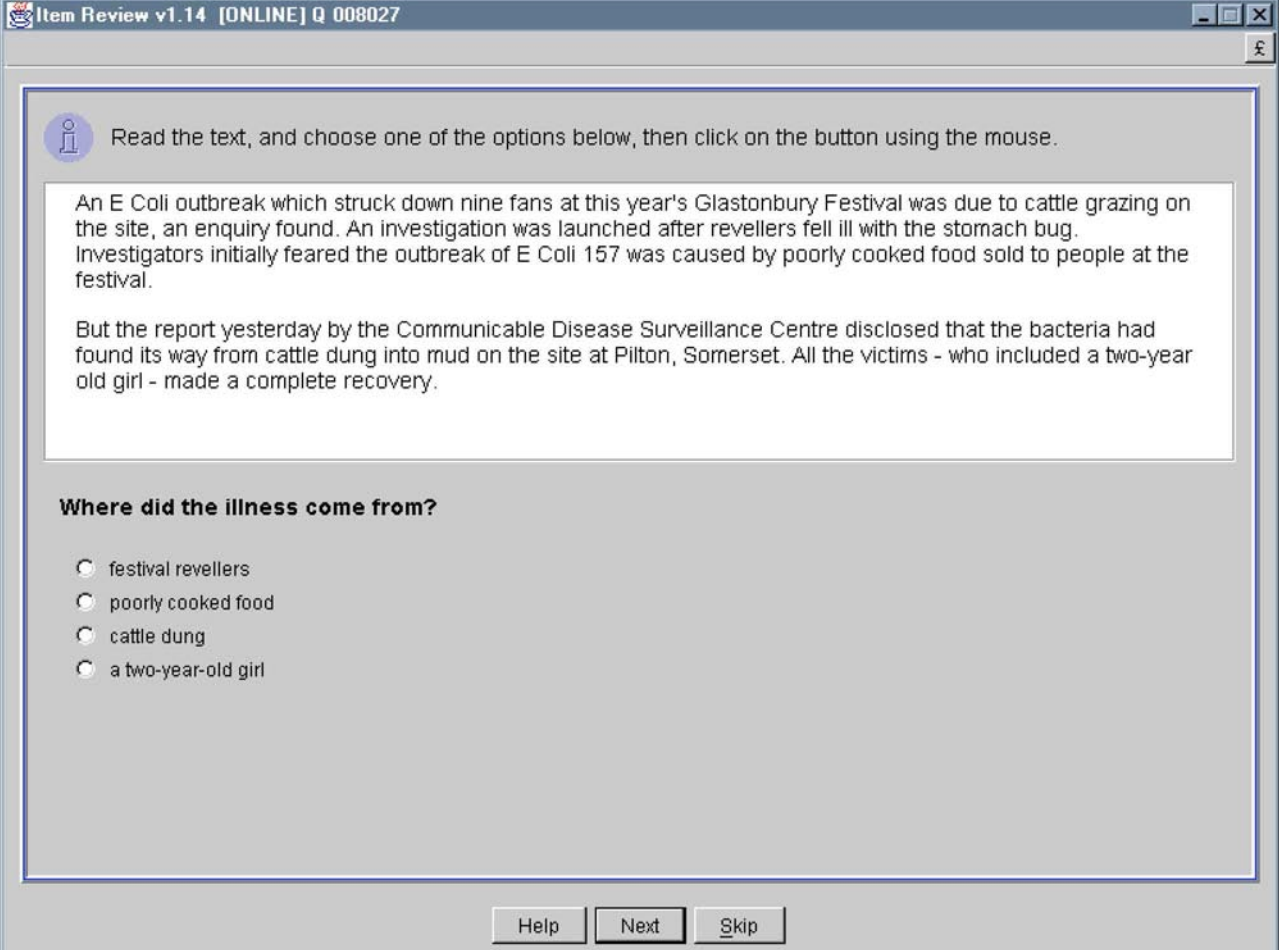
Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
3	178	.815	.556	.579
4	187	.904	.521	.555

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
120 R008026	4	-.257	.045	1.269	3	.736	-.491	-.637	-1.042

DIALANG

B1

Item Review v1.14 [ONLINE] Q 008027

ℹ️ Read the text, and choose one of the options below, then click on the button using the mouse.

An E Coli outbreak which struck down nine fans at this year's Glastonbury Festival was due to cattle grazing on the site, an enquiry found. An investigation was launched after revellers fell ill with the stomach bug. Investigators initially feared the outbreak of E Coli 157 was caused by poorly cooked food sold to people at the festival.

But the report yesterday by the Communicable Disease Surveillance Centre disclosed that the bacteria had found its way from cattle dung into mud on the site at Pilton, Somerset. All the victims - who included a two-year old girl - made a complete recovery.

Where did the illness come from?

○ festival revellers

○ poorly cooked food

○ cattle dung

○ a two-year-old girl

Help   Next   Skip

DIALANG

Content analysis

Text Characteristics:

Task:	8027
1. Text Source:	Newspapers
2. Authenticity:	Authentic
3. Discourse type:	Mainly Narrative
4. Domain:	Public
5. Topic:	7. Health and bodycare
6. Nature of Content:	mostly concrete content
7. Text Length:	101
8. Vocabulary:	rather extended
9. Grammar:	Limited range of complex structures
Comprehensible by learner at CEF level:	B1/B2

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Recognise and Retrieve
15. Operations:	Explicit
15. Operations:	Conclusion
Item Level Estimated:	B1/B2
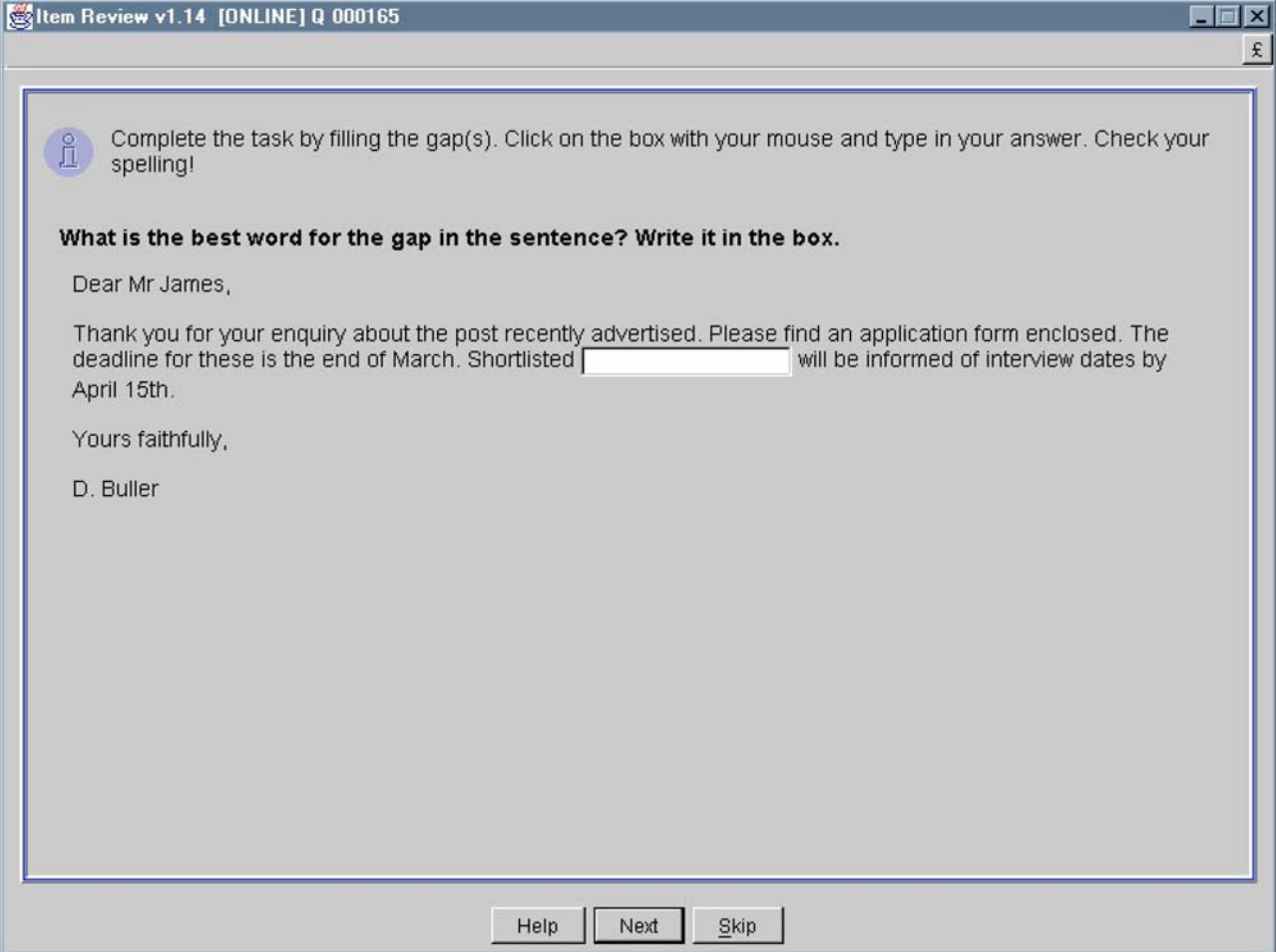
Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
1	183	.645	.511	.502
2	171	.743	.450	.452

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
85 R008027	3	.038	.044	7.771	6	.255	.535	.143	-.601

DIALANG

B2

Item Review v1.14 [ONLINE] Q 000165

£

Complete the task by filling the gap(s). Click on the box with your mouse and type in your answer. Check your spelling!

What is the best word for the gap in the sentence? Write it in the box.

Dear Mr James,

Thank you for your enquiry about the post recently advertised. Please find an application form enclosed. The deadline for these is the end of March. Shortlisted [   ] will be informed of interview dates by April 15th.

Yours faithfully,

D. Buller

Help   Next   Skip

Content analysis
Text Characteristics:

Task:	165
1. Text Source:	Business letter
2. Authenticity:	Authentic
3. Discourse type:	Mainly Instructive
4. Domain:	Occupational
5. Topic:	15. Other
Specify other topic:	employment
6. Nature of Content:	mostly concrete content
7. Text Length:	42
8. Vocabulary:	rather extended
9. Grammar:	mainly simple structures
Comprehensible by learner at CEF level:	B1

Item Characteristics:

Item:	1
14. Item Type:	9. Gap filling (one word)
15. Operations:	Make inferences
15. Operations:	Implicit
15. Operations:	Detail
Item Level Estimated:	B2

Statistical analysis
Information from piloting:
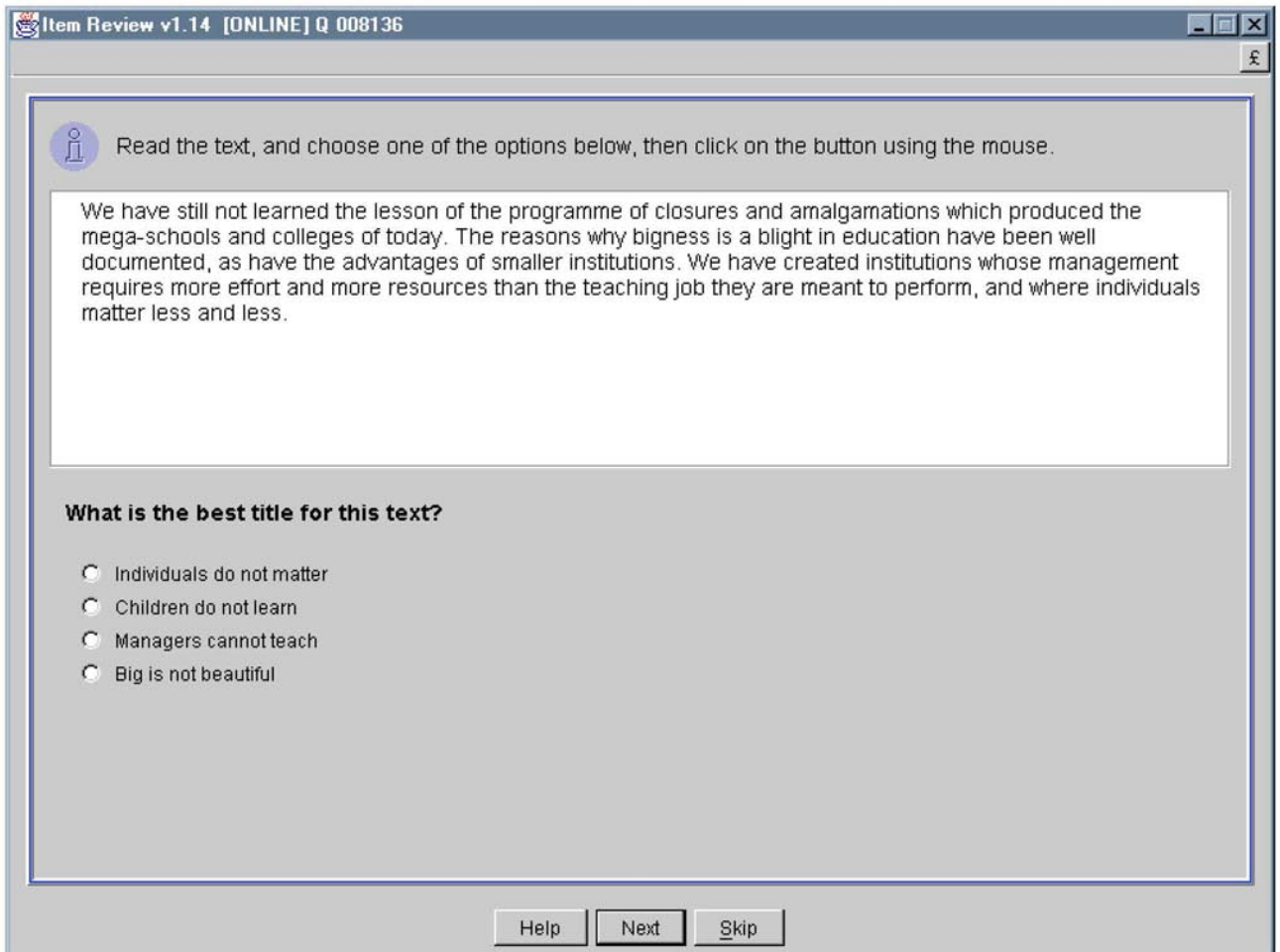
Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
3	178	.376	.514	.535
4	187	.294	.503	.515

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
117 R000165	4	.671	.037	3.372	5	.643	.837	.539	.265

B2

Item Review v1.14 [ONLINE] Q 008136

Read the text, and choose one of the options below, then click on the button using the mouse.

We have still not learned the lesson of the programme of closures and amalgamations which produced the mega-schools and colleges of today. The reasons why bigness is a blight in education have been well documented, as have the advantages of smaller institutions. We have created institutions whose management requires more effort and more resources than the teaching job they are meant to perform, and where individuals matter less and less.

What is the best title for this text?

○ Individuals do not matter

○ Children do not learn

○ Managers cannot teach

○ Big is not beautiful

Help   Next   Skip

Content analysis

Text Characteristics:

Task:	8136
1. Text Source:	Magazines
2. Authenticity:	Authentic
3. Discourse type:	Mainly Argumentative
4. Domain:	Public
4. Domain:	Educational
5. Topic:	8. Education
6. Nature of Content:	mainly abstract content
7. Text Length:	71
8. Vocabulary:	rather extended
9. Grammar:	Limited range of complex structures
Comprehensible by learner at CEF level:	B2/C1

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Evaluate
15. Operations:	Explicit
15. Operations:	Main idea/gist
Item Level Estimated:	B2

Statistical analysis
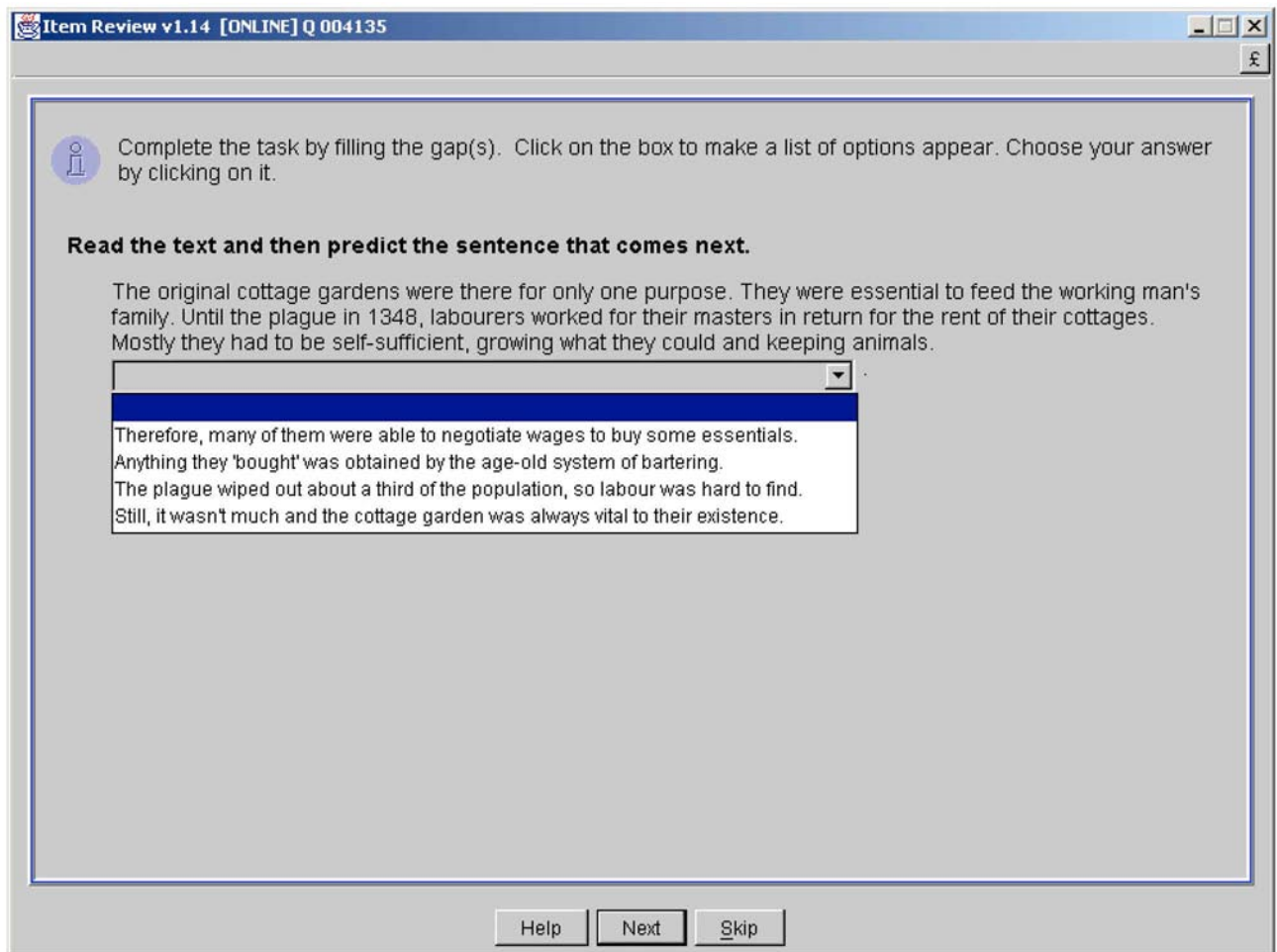
Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
2	171	.550	.521	.514
3	178	.528	.485	.478

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
103 R008136	3	.349	.043	8.853	6	.182	.620	.692	.141

DIALANG

C1

Item Review v1.14 [ONLINE] Q 004135

Complete the task by filling the gap(s). Click on the box to make a list of options appear. Choose your answer by clicking on it.

Read the text and then predict the sentence that comes next.

The original cottage gardens were there for only one purpose. They were essential to feed the working man's family. Until the plague in 1348, labourers worked for their masters in return for the rent of their cottages. Mostly they had to be self-sufficient, growing what they could and keeping animals.

Therefore, many of them were able to negotiate wages to buy some essentials.
Anything they 'bought' was obtained by the age-old system of bartering.
The plague wiped out about a third of the population, so labour was hard to find.
Still, it wasn't much and the cottage garden was always vital to their existence.

Help   Next   Skip

Content analysis

Text Characteristics:

Task:	4135
1. Text Source:	Reference books
2. Authenticity:	Authentic
3. Discourse type:	Mainly Expository
4. Domain:	Public
5. Topic:	15. Other
Specify other topic:	History
6. Nature of Content:	fairly extensive abstract content
7. Text Length:	51
8. Vocabulary:	rather extended
9. Grammar:	Limited range of complex structures
Comprehensible by learner at CEF level:	B2

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Make inferences
15. Operations:	Implicit
15. Operations:	Conclusion
15. Operations:	Text Structure/Connections between parts
Item Level Estimated:	B2/C1

Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
1	183	.164	.258	.203
4	187	.203	.270	.235

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
66 R004135	2	1.226	.072	7.375	6	.288	-.505	1.394	1.606

◗DIALANG

Dialang English Listening Comprehension Items

General information

There were 276 – 330 pilot test takers per each listening item. Overall data – model fit was satisfactory (p = .24).

The cut-offs on the theta scale for the CoE levels are based on standard setting procedures where 9 – 11 experts judged the items (# of experts varied from skill to skill); the standard setting data were analysed with a special programme designed by Norman Verhelst at CITO.

Cut-off points for the Framework levels for the DIALANG English listening items:

under A1: theta under -.90
A1: theta between -.90 and -.50
A2: theta between -.50 and +.06
B1: theta between +.06 and +.56
B2: theta between +.56 and +1.00
C1: theta between +1.00 and +1.42
C2: theta above +1.42

A2

Content analysis

Text Characteristics:

Task:	20851
1. Text Source:	News broadcasts
2. Authenticity:	Authentic
3. Discourse type:	Mainly Narrative
4. Domain:	Public
5. Topic:	3. Daily life
6. Nature of Content:	mostly concrete content
7. Text Length:	12secs
8. Vocabulary:	mostly frequent vocabulary
9. Grammar:	mainly simple structures
10. Text Speed:	normal
11. No of participants:	one
12. Accent Standard:	Standard pronunciation
13. Clarity of Articulation:	normally articulated
14. How often played:	played once
Comprehensible by learner at CEF level:	B1

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Make inferences
15. Operations:	Explicit
15. Operations:	Detail
Item Level Estimated:	B1

Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
9	162	.877	.643	.657
12	141	.894	.473	.493

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
4 L020851	4	-.253	.056	.714	2	.700	.572	-.495	-.739

A2

Content analysis

Text Characteristics:

Task:	4276
1. Text Source:	Public announcements and instructions
2. Authenticity:	Authentic
3. Discourse type:	Mainly Instructive
4. Domain:	Public
5. Topic:	5.Travel
6. Nature of Content:	only concrete content
7. Text Length:	32
8. Vocabulary:	mostly frequent vocabulary
9. Grammar:	mainly simple structures
10. Text Speed:	slow
11. No of participants:	one
12. Accent Standard:	Standard pronunciation
13. Clarity of Articulation:	normally articulated
14. How often played:	played once
Comprehensible by learner at CEF level:	A2/B1

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Recognise and Retrieve
15. Operations:	Explicit
15. Operations:	Detail
Item Level Estimated:	A2

Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
10	168	.899	.546	.575
11	135	.904	.553	.589

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
35 L004276	5	-.361	.049	.849	1	.357	-.860	-.860	-.084

A2

Content analysis

Text Characteristics:

Task:	7616
1. Text Source:	Interpersonal dialogues and conversations
2. Authenticity:	Authentic
3. Discourse type:	Mainly Argumentative
4. Domain:	Personal
5. Topic:	3. Daily life
6. Nature of Content:	only concrete content
7. Text Length:	2 secs
8. Vocabulary:	only frequent vocabulary
9. Grammar:	only simple structures
10. Text Speed:	normal
11. No of participants:	two
12. Accent Standard:	Standard pronunciation
13. Clarity of Articulation:	normally articulated
14. How often played:	played once
Comprehensible by learner at CEF level:	A2

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Make inferences
15. Operations:	Implicit
15. Operations:	Main idea/gist
Item Level Estimated:	A2

Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
10	168	.887	.617	.645
11	135	.941	.552	.599

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
39 L007616	5	-.395	.050	.453	1	.501	-.903	-.787	-1.351

A2

Content analysis

Text Characteristics:

Task:	5480
1. Text Source:	Interpersonal dialogues and conversations
2. Authenticity:	Authentic
3. Discourse type:	Mainly Instructive
4. Domain:	Personal
4. Domain:	Public
5. Topic:	5.Travel
6. Nature of Content:	only concrete content
7. Text Length:	28secs
8. Vocabulary:	only frequent vocabulary
9. Grammar:	only simple structures
10. Text Speed:	slow
11. No of participants:	two
12. Accent Standard:	Strong regional accent
13. Clarity of Articulation:	clearly articulated
14. How often played:	played once
Comprehensible by learner at CEF level:	A2

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Recognise and Retrieve
15. Operations:	Explicit
15. Operations:	Main idea/gist
Item Level Estimated:	A2

Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
11	135	.852	.584	.608
12	141	.922	.578	.616

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
56 L005480	5	-.258	.051	.322	1	.570	.040	-.896	-.893

B1

Content analysis

Text Characteristics:

Task:	7615
1. Text Source:	Interpersonal dialogues and conversations
2. Authenticity:	Pedagogic
3. Discourse type:	Mainly Descriptive
4. Domain:	Personal
5. Topic:	10. Food and drink
6. Nature of Content:	only concrete content
7. Text Length:	24
8. Vocabulary:	mostly frequent vocabulary
9. Grammar:	only simple structures
10. Text Speed:	normal
11. No of participants:	one
12. Accent Standard:	Standard pronunciation
13. Clarity of Articulation:	normally articulated
14. How often played:	played once
Comprehensible by learner at CEF level:	A2/B1

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Make inferences
15. Operations:	Explicit
15. Operations:	Detail
Item Level Estimated:	B1

Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
9	162	.660	.545	.542
10	168	.548	.532	.505

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
24 L007615	3	.301	.046	9.319	6	.156	-.277	-1.120	-.416

B1

Content analysis

Text Characteristics:

Task:	21135
1. Text Source:	TV/ radio documentaries
2. Authenticity:	Authentic
3. Discourse type:	Mainly Narrative
4. Domain:	Public
5. Topic:	5.Travel
6. Nature of Content:	fairly extensive abstract content
7. Text Length:	52secs
8. Vocabulary:	rather extended
9. Grammar:	Limited range of complex structures
10. Text Speed:	normal
11. No of participants:	one
12. Accent Standard:	Standard pronunciation
13. Clarity of Articulation:	clearly articulated
14. How often played:	played once
Comprehensible by learner at CEF level:	B2

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Make inferences
15. Operations:	Explicit
15. Operations:	Main idea/gist
Item Level Estimated:	B2

Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
9	162	.630	.478	.471
10	168	.518	.487	.465

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
26 L021135	3	.361	.046	13.931	6	.030	1.106	.514	1.056

DIALANG

B1

Content analysis

Text Characteristics:

Task:	20849
1. Text Source:	News broadcasts
2. Authenticity:	Authentic
3. Discourse type:	Mainly Narrative
4. Domain:	Public
5. Topic:	15. Other
Specify other topic:	Crime
6. Nature of Content:	mostly concrete content
7. Text Length:	15secs
8. Vocabulary:	rather extended
9. Grammar:	Limited range of complex structures
10. Text Speed:	normal
11. No of participants:	one
12. Accent Standard:	Standard pronunciation
13. Clarity of Articulation:	normally articulated
14. How often played:	played once
Comprehensible by learner at CEF level:	B1/B2

Item Characteristics:

Item:	1
14. Item Type:	1. Multiple choice
15. Operations:	Recognise and Retrieve
15. Operations:	Explicit
15. Operations:	Main idea/gist
Item Level Estimated:	B1

Statistical analysis

Information from piloting:

Booklet number	n size	Facility	Rpbis (unweighted)	Rpbis (weighted)
11	135	.659	.708	.742
12	141	.695	.604	.644

Results of calibration:

nr label	A	B	SE(B)	S	DF	P	M	M2	M3
53 L020849	5	.218	.037	5.385	4	.250	-.305	.044	-.920