

# OBJECTIVE TESTING SECTION

UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS

**Examples of Statistics  
and Questions for  
Multiple Choice**





# Contents

## 4 Item Statistics

*A Brief Explanation of Statistics*  
*Facility*  
*Discrimination*  
*Typical GCE A Level Statistics*

## 5 Judging Items on the Basis of Statistics

*Low Facility & Discrimination*  
*Positive Distractors*  
*Low Discrimination*

## 6 Item Analysis Example 1

*Human & Social Biology*

## 7 Item Analysis Example 2

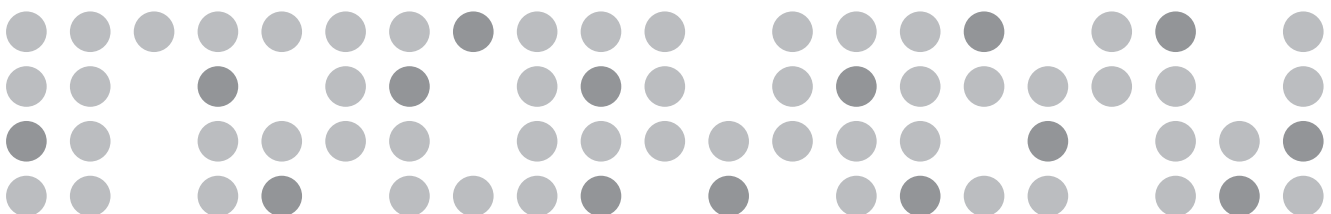
*Human & Social Biology*

## 8 Item Analysis Example 3

*Human & Social Biology*

## 9 Item Analysis Example 4

*Geography*



## ITEM STATISTICS

The responses of the candidates to a multiple choice test are analysed to produce statistics relating to the individual items and overall test statistics.

### Brief Explanation of Statistics

The most useful statistics calculated are an item's facility and discrimination.

#### Facility

Facility is the proportion of candidates choosing the key option. Easy items have high facilities. Ideally, each distractor should attract some candidates, even if the proportion of all candidates who are distracted is small. In tests which aim for a mean score of around 60–65%, all distractors should attract at least 5% of the candidates.

#### Discrimination

Discrimination measures an item's effectiveness at separating able and less able candidates. The relationship between choice of option and performance on the whole test is calculated. Many correlation coefficients exist which could do this. The point biserial selected as one of the variables is dichotomous (e.g. right / wrong on an item) and the other is relatively finely scaled (e.g. total score on a test).

The Discrimination value has a theoretical range from  $-1.0$  to  $+1.0$ . Values greater than zero indicate a positive relationship between choice of option and overall success on the test. Values below zero indicate a negative relationship. Values above 0.25 for a key option indicate that an item is discriminating adequately. A value of zero implies that there was no relationship – candidates' choice is random.

### Typical GCE A Level Item Statistics

The facility and discrimination limits for a 4-option GCE A Level test are:

- facility: 0.25 to 0.80
- discrimination: 0.25 or higher

The lower facility figure represents the 'guessing level', the facility which is expected if all candidates guess at random.

The table is a sample item analysis of the results of four GCE A level items. The mean score on the test was around 60 %.

item number	value	Options (* indicates key)				no response
		A	B	C	D	
1	facility	0.174	0.058	0.691	0.071	0.006
	discrimination	-0.006	-0.187	0.344	-0.148	-0.149
2	facility	0.085	0.599*	0.257	0.039	0.020
	discrimination	-0.163	0.275*	-0.127	-0.111	-0.087
3	facility	0.061	0.206*	0.092	0.614*	0.026
	discrimination	-0.221	-0.053*	-0.223	0.330*	-0.139
4	facility	0.054	0.121	0.131	0.663*	0.026
	discrimination	-0.124	-0.088	-0.260	0.337*	-0.139

# JUDGING ITEMS ON THE BASIS OF STATISTICS

## **Low Facility and Discrimination**

A valid item, which is appropriate as a measure of the achievement of properly prepared candidates, may produce statistics outside the limits because the sample of candidates for that test were ill-prepared.

Bearing this in mind, certain items may be included in a test as 'curriculum development' items to introduce novel ideas into the teaching of the subject, even though the items may be relatively unsuccessful at discriminating between candidates or at helping them to show positive achievement. As the inclusion of such items reduces the reliability of the test, no more than one or two should be included.

## **Positive Distractors**

Distractors with positive discrimination values may show that a significant number of able candidates opted for these distractors. Items with positive distractors should be inspected closely with a view to possible revision, as the item may contain an ambiguity or other item writing fault.

As a rule of thumb, items with distractors producing discrimination values of more than +0.1 should be discarded or revised, unless the positive distractor represents a commonly held misconception which an examiner wishes to bring to the attention of teachers. Where positive distractors have a discrimination value of below 0.1, however, the corresponding facility should be taken into account (and the number of candidates in the sample). If only a small proportion of candidates is involved, a positive discrimination value may not be significant.

## **Low Discrimination**

This may indicate that the item is a poor measure of whatever it is measuring but it may be measuring something different from the other items in the test.

## ITEM ANALYSIS EXAMPLES

The four example items are taken from tests where the emphasis is on positive achievement, i.e. where high facility values and high mean scores are expected. Items 1–3 are from the Human and Social Biology bank, item 4 is from a Geography bank.

There is something wrong with the statistics of all the items. Suggested solutions are provided in a box at the end of each item.

### EXAMPLE 1

1 Why is it important to keep a cooking area clean?

- A to prevent rats getting in
- B to stop smells getting in
- C to keep food cool
- D to discourage houseflies

#### Statistics

value	A	B	C	D	no response
facility	0.061	0.012	0.008	0.915*	0.004
discrimination	-0.081	-0.119	-0.133	0.165*	-0.031

**Problems** – Low discrimination. 0.165 is below the lower design limit of 0.2.

Option **C** took only 0.008 of candidates, less than 1%, a marginally effective distractor although one such distractor may be expected when the facility is as high as 0.915.

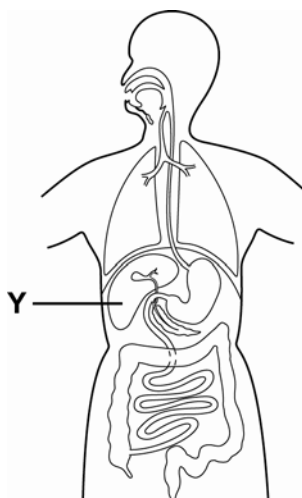
#### Suggested revision

1	Houseflies will be discouraged from settling in the cooking area if it is kept
A	clean.
B	humid.
C	warm.
D	well-lit.

**Comments** – Transposition of stem and key option is one of the first strategies to be attempted in trying to rescue poor items.

## EXAMPLE 2

2 The diagram shows the positions of some of the main organs in the body.



What is organ Y?

- A brain
- B liver
- C lungs
- D pancreas

### Statistics

value	A	B	C	D	no response
facility	0.005	0.925*	0.029	0.041	0.001
discrimination	-0.046	0.311*	-0.203	-0.225	-0.012

**Problems** – Option A is virtually inoperable. Only half of one percent of candidates opted for it. Even in tests designed to show positive achievement, such options should be scrutinised.

### Suggested revision

2 The diagram shows the positions of the main organs in the body.

(diagram as before)

What is organ Y?

- A liver
- B lungs
- C pancreas
- D stomach

**Comments** – Clearly, any number of organs on the syllabus could be tried out. Improvement is straightforward.

### EXAMPLE 3

3 What is the most important reason for sleeping in a well-ventilated room?

- A The air is humid and warm.
- B The air is humid and cool.
- C The air carries away harmful germs.
- D The air keeps heat in.

#### Statistics

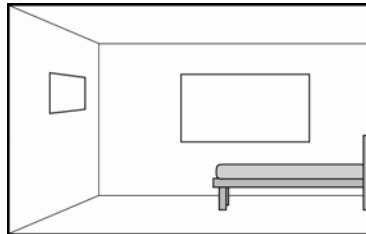
value	A	B	C	D	no response
facility	0.062	0.482	0.424*	0.030	0.002
discrimination	-0.123	-0.019	0.138*	-0.162	-0.020

**Problems** – low discrimination of 0.138

Option **B** is more popular than the key option and the very small negative correlation with the total test score (-0.019) suggests that some of the more able candidates may have been distracted by it. (This may be understood when it is explained that the statistics relate to Singaporean candidates.)

#### Suggested revision

3 The diagram shows a bedroom with a large window and a small window.



For the most healthy sleeping conditions, how should the windows be arranged?

	small window	large window
A	closed	closed
B	closed	opened
C	opened	closed
D	opened	opened

**Comments** – This new item provided an example of the central problem with which the item is concerned and allows a significant range of focus. The item is now more clearly concerned with ventilation. The specifications of this revised item may vary from that of the original, but this represents a way of saving the original idea.

**EXAMPLE 4**

4 A volcano is described as dormant when

- A it is erupting.
- B it is likely to erupt again.
- C people think it will not erupt again.
- D it has a large crater.

value	A	B	C	D	no response
facility	0.061	0.012	0.008	0.915*	0.004
discrimination	-0.081	-0.119	-0.133	0.165*	-0.031

**Problem** – A low facility of 0.420 is below the lower design limit of 0.5.

**Suggested revision**

4 The world map shows four shaded areas.

In which area is there **most** likely to be a dormant volcano?

