

What makes GCSE examination questions difficult? Outcomes of manipulating difficulty of GCSE questions.

Hannah Fisher-Hoch, Sarah Hughes and Tom Bramley

Research and Evaluation

University of Cambridge Local Examinations Syndicate

Paper presented at the British Educational Research Association Annual Conference
(September 11-14 1997: University of York)

ABSTRACT

Candidates' scripts from GCSE Mathematics, Science, Geography and O level English examination papers have been analysed for common errors. The possible source of difficulty (SOD) for each of these errors has been hypothesised. These SODs have been used to inform the process of manipulating the questions which have then been trialled. The results of the trials have enabled us to examine the effects of changes in wording, layout, diagrams etc. with a view to informing examiners.

INTRODUCTION

The question difficulty project (QDP) is looking at the question "What makes one examination question more difficult than another?". If a significant number of candidates fail to get the correct response to a particular examination question, then that question is considered to be difficult. This lack of success may be the result of being unable to recall the knowledge or strategy required to answer the question: 'valid difficulty'.

Alternatively, the lack of success may be the result of being unable to access the question for reasons unrelated to the subject matter such as language, incoherent diagram etc.: 'invalid difficulty'. The aim of the project is to identify Sources of Difficulty (SODs), both valid and invalid, with a view to informing the examination writing process.

The QDP is investigating five subjects, taken by 16 year olds (Year 11), to identify SODs in examination questions. The subjects under investigation are: GCSE Mathematics, Geography and Science, IGCSE French and O level English.

The syllabus (see Fig.1 below) for each subject was chosen on the basis that the examinations consisted of short written response type questions. These questions give an optimum amount of information for the qualitative analysis of the candidates responses.

Fig. 1 : Syllabus chosen for each subject.

Subject	Mathematics SMP 11-16	Geography Syllabus A	Science Salter's	English	French
	GCSE	GCSE	GCSE	O level	IGCSE
Syllabus No.	1663	1575	1774	1120/3/4/5	0520
Year	June 1994	June 1994	June 1995	Nov. 1994	June 1995

The identification of SODs needs to be as precise as possible and for each of the papers the questions were broken down into 'bits', the smallest unit of a question for which marks can be awarded. Using statistical analysis the most difficult question bits for each subject were identified. These question bits were then subjected to an error analysis, where the candidates' scripts were scrutinised for common errors. This method of error analysis was based on the methods of Pollitt et al (1985) and Griffith and McLone (1989). The identification of common errors led to a list of hypothesised SODs for each of the subjects (Appendix I, page 14). The list of SODs for each subject has, until now, remained independent. However, overlapping of SODS across the subjects can already be seen, and this will enable us to make generalisations.

The SODs have been used to inform the manipulation of the original questions to make them easier or more difficult. The resulting questions have been trialled on Year 11 students to investigate the effect on performance.

At BERA 96 the team reported on the methodology and initial findings from the analysis of the Mathematics examination papers. Over the last few months the project has moved forward: analysing the Geography, English and Science examination papers; completing an initial analysis of French; trialling papers containing manipulated questions in Mathematics, Geography, Science and English and analysing the results of the Mathematics trials. This paper will report on the construction of trial papers for Mathematics, Geography, Science and English and the outcomes of the Mathematics trials.

MANIPULATIONS

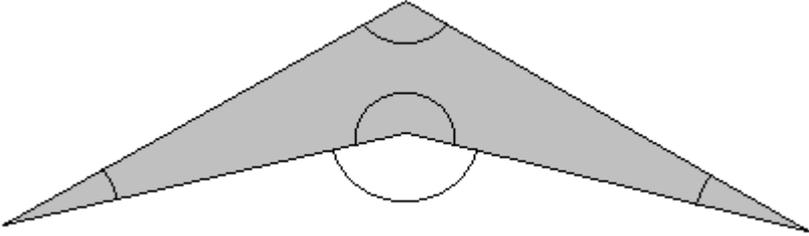
The identified SODs were used to inform the manipulation of questions for trialling in schools. The questions chosen for manipulation were selected on the basis that they would represent as wide a sample of SODs as possible. Trial papers were designed to be taken during a school period, usually sixty minutes. This time constraint meant that there was a limit to the number of questions that could be trialled. However, most of the identified SODs were included. The timing was calculated on a pro rata basis from the length of the original papers.

The original questions from which the SODs were identified were manipulated. Original question bits were manipulated in order to avoid creating further SODs by writing new questions.

The initial idea was to manipulate only the SOD that had been identified, but it was rarely possible to manipulate a question so precisely. In practice other aspects of the question were also changed in order to accommodate the manipulation of the SOD. For example, the question in Figs. 2 and 3 show the original and a manipulated version of a question from Paper 1 of the Mathematics syllabus.

Fig. 2 : Paper 1 : Question 19 : Original question

19	This shape is called an arrow head.
	Mark and label clearly
(a)	an acute angle, [1]
(b)	a reflex angle. [1]

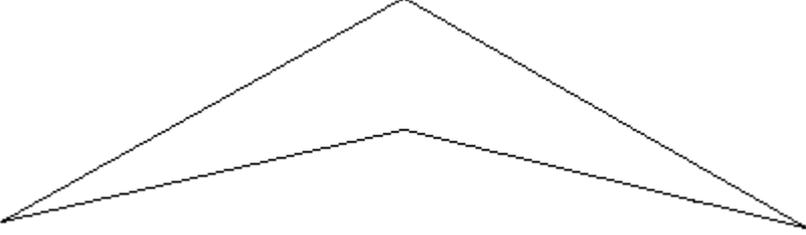


There were various SODs associated with this question most notably those relating to the diagram ('ambiguous resources'). The candidates had problems when they were asked to 'mark and label' angles that have already been marked. One of the versions that was created for this question was to remove the markings from the diagram as shown in Fig. 3.

Fig. 3 : Paper 1 : Question 19 : Manipulated version

19 This shape is called an arrow head.
Mark and label **clearly**

(a) an acute angle, [1]
(b) a reflex angle. [1]



The SOD was addressed by removing the shading and the marking to minimise the confusion created in the original question. However, these changes in turn increased the choice of angles that could be marked by the candidates, which created an unwanted change in the question. This was a recurrent problem through many of the questions. In order to monitor these effects, further qualitative analysis was carried out while marking the trial scripts.

CONSTRUCTION OF TRIAL PAPERS

The trial papers were constructed from the original examination papers analysed for SODs. Each of the GCSE trial subjects had tiered papers corresponding to Basic, Central and Further Tiers.

The trial papers for each subject were constructed slightly differently, as detailed below. On advice from examiners in the subjects, whole questions or the question up to the bit being manipulated were maintained. This was in order to avoid creating further difficulties by changing any part of the question preceding the bit under investigation. However, in some cases the bits subsequent to those which had been manipulated were omitted.

The papers were designed to look like examination papers with similar layout to the original papers. The students sat the papers under test conditions and had, in most cases, been asked to revise for the test.

Mathematics manipulations

The original examination consisted of six papers increasing in difficulty from one to six. Each candidate took two consecutive papers. There was an overlap of papers across the Tiers of entry. For some of the questions to be manipulated there were innumerable variations that could be constructed and for others only one or two. For practical purposes there were either two or four versions, including the original, selected for use in the papers. The trial papers consisted of three Tiers: X, Y and Z, containing questions from papers 4, 5 and 6; 3 and 4; 1, 2 and 3 respectively. There were four papers created at Basic and Further Tier and eight papers at Central Tier. The naming of the papers is shown in Fig. 4.

Fig. 4 : Naming of Mathematics trial papers

Tier		
<i>Basic</i>	<i>Central</i>	<i>Further</i>
ZA, ZB, ZC, ZD	YA, YB, YC, YD, YAB, YBC, YCD, YDA	XA, XB, XC, XD

The eight papers at Tier Y consisted of four papers (YAB, YBC, YCD, YDA) which consisted of a combination of the original four (YA, YB, YC, YD) to help with statistical equating of data. In order to be able to compare difficulty across the Tiers the same questions from Papers 3 and 4 appeared on both Tiers.

Geography manipulations

The original examination consisted of three papers increasing in difficulty from one to three. The papers were constructed so that the same topics were covered on all Tiers but the questions became more structured. Candidates sat one of the three papers. Geography did not lend itself to as many variations as Mathematics and therefore, for each question, two further versions were created. The trial papers consisted of three Tiers: 1, 2 and 3, each containing three papers: a, b and c (see Fig. 5). In order to be able to compare difficulty there was an overlap of questions across Tiers and Papers within Tiers.

Fig. 5 : Naming of Geography papers

Tier		
<i>Basic</i>	<i>Central</i>	<i>Further</i>
1a, 1b, 1c	2a, 2b, 2c	3a, 3b, 3c

English manipulations

The original examination consisted of two papers taken by all candidates. One of the papers was a comprehension paper and the other a composition paper. Only the comprehension paper was investigated for this study. The paper started with a passage followed by short answer questions and then a summary exercise. The length of the paper meant that for the purposes of trialling the paper was split in half, one half containing the short answer questions and the other the summary. This resulted in five versions, including the original, for both the short answer paper and the summary paper. There were common questions across the papers using some short answer questions on the ‘summary’ papers. As O level English is not tiered, the papers were differentiated by colour.

Science manipulations

The original examination consisted of six papers, each Tier consisting of a pair of papers: 1 and 2 (Basic Tier); 3 and 4 (Central Tier); 5 and 6 (Further Tier). Papers 1, 3 and 5 and 2, 4 and 6 increase in difficulty and have the same content. There are overlapping questions between Basic and Central Tier and Central and Further Tier. The questions that overlap the Tiers are identical, but may have extra parts attached for the higher Tier. As with Geography there were only three versions, including the original, of each Science question. The trial consisted of three papers; A, B and C at each of three Tiers X, Y and Z (see Fig. 6). There was an overlap of questions across and within Tiers.

Fig. 6 : Naming of Science papers

Tier		
<i>Basic</i>	<i>Central</i>	<i>Further</i>
XA, XB, XC	YA, YB, YC	ZA, ZB, ZC

TRIALLING OF MATHEMATICS PAPERS

Each of the sixteen Mathematics papers was taken by approximately 50 students, from 5 schools. The statistical analysis of the resulting data gave facility values for each question bit. The changes in these values, between the original and the manipulated versions, enabled us to analyse the comparative performance of different versions of the question bits.

Results

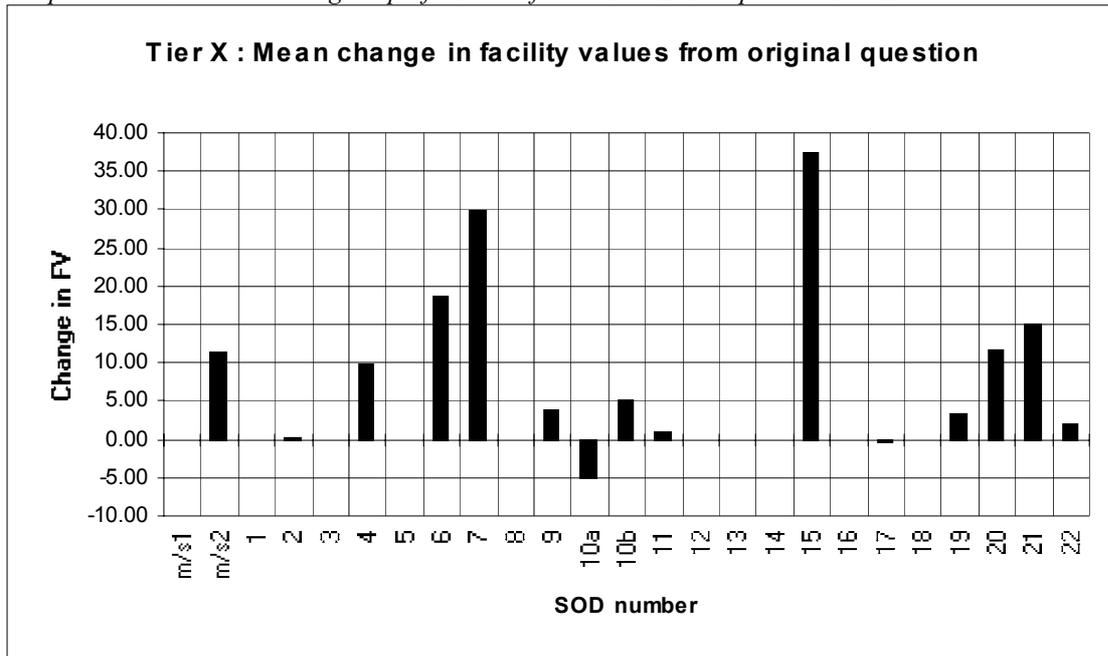
For each question bit the SOD was identified and manipulated and the associated change in facility value calculated. The table (Fig. 7) and the graphs on the next pages show the mean difference in performance, for each Tier and overall, of the SODs. It can be seen from Fig. 7 that for most of the manipulations the performance was improved.

Fig. 7 : Table showing the mean change in Facility Value for each Source of Difficulty manipulated in the Mathematics trials.

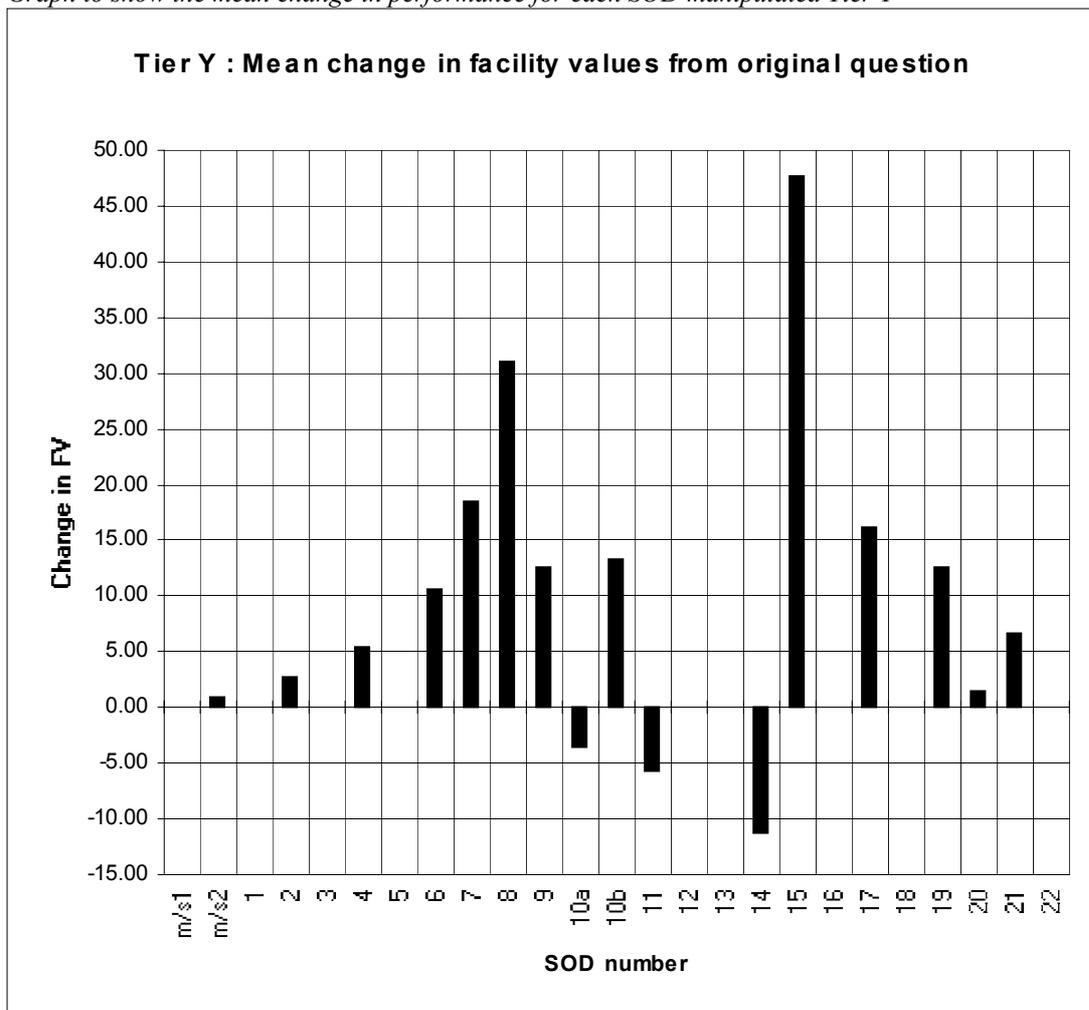
SOD		Mean change in F.V. ¹			
		Tier Z	Tier Y	Tier X	All Tiers
m/s1	mark scheme (collapsed)	3.32	0.00	n/a	2.77
m/s2	mark scheme (allocation)	16.40	1.00	11.40	10.32
1	command words	n/a	n/a	n/a	n/a
2	mathematical language	-1.10	2.70	0.13	0.58
3	mathematics v. everyday language	n/a	n/a	n/a	n/a
4	technical notation	4.77	5.33	9.72	6.78
5	non-mathematical language	n/a	n/a	n/a	n/a
6	recall strategy	9.67	10.60	18.65	11.93
7	recall knowledge	9.18	18.50	29.87	17.05
8	number of steps	15.50	31.00	n/a	23.25
9	dense presentation	-6.30	12.55	3.80	5.28
10a	context (data handling)	-15.40	-3.53	-5.14	-4.70
10b	context (other)	4.39	13.36	5.09	6.98
11	stated principle	n/a	-5.7	0.93	-0.77
12	combination of topics	n/a	n/a	n/a	n/a
13	isolated skill	n/a	n/a	n/a	n/a
14	mathematical sequencing	n/a	-11.20	n/a	-14.00
15	arithmetic errors	n/a	47.60	37.50	42.55
16	alternative strategies	n/a	n/a	n/a	n/a
17	abstraction required	12.03	16.10	-0.30	10.38
18	spatial representation required	-3.97	n/a	n/a	-3.97
19	paper layout	-6.30	12.55	3.33	4.59
20	highlighting	-2.85	1.40	11.55	3.76
21	ambiguous resources	8.49	6.60	15.10	8.64
22	irrelevant information	n/a	0.00	2.10	1.05

¹ For all the questions where any given SOD was manipulated the numerical change between the performance for the manipulation and the original version (at trialling) was calculated. The mean value of these changes represents the mean change in F.V.

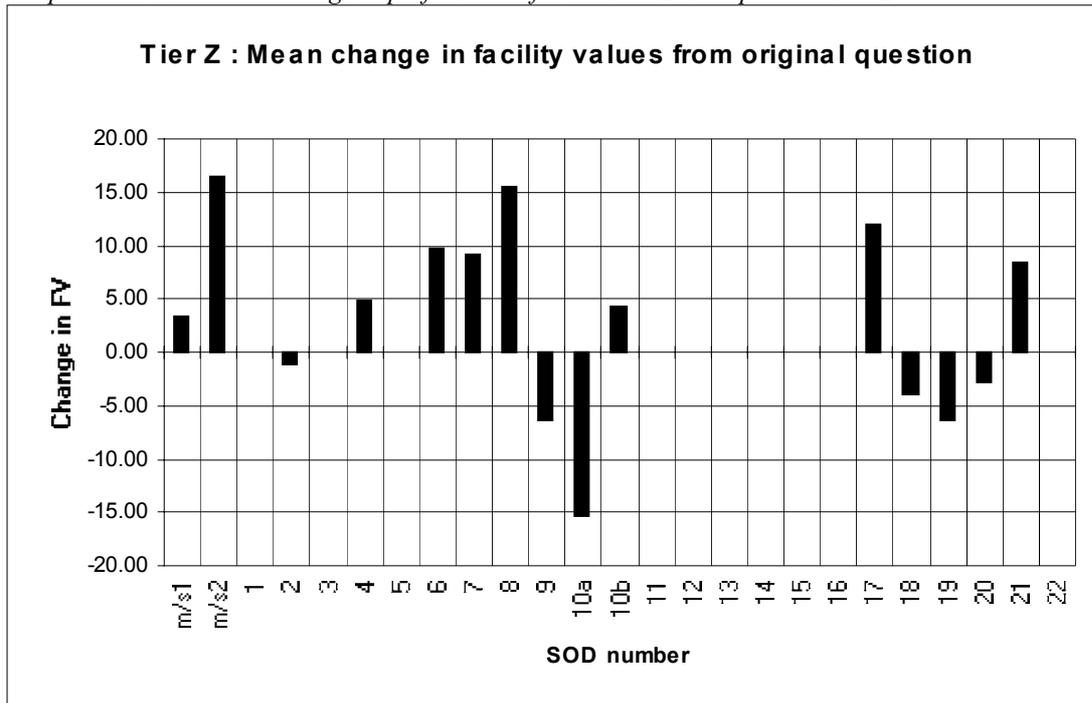
Graph to show the mean change in performance for each SOD manipulated Tier X



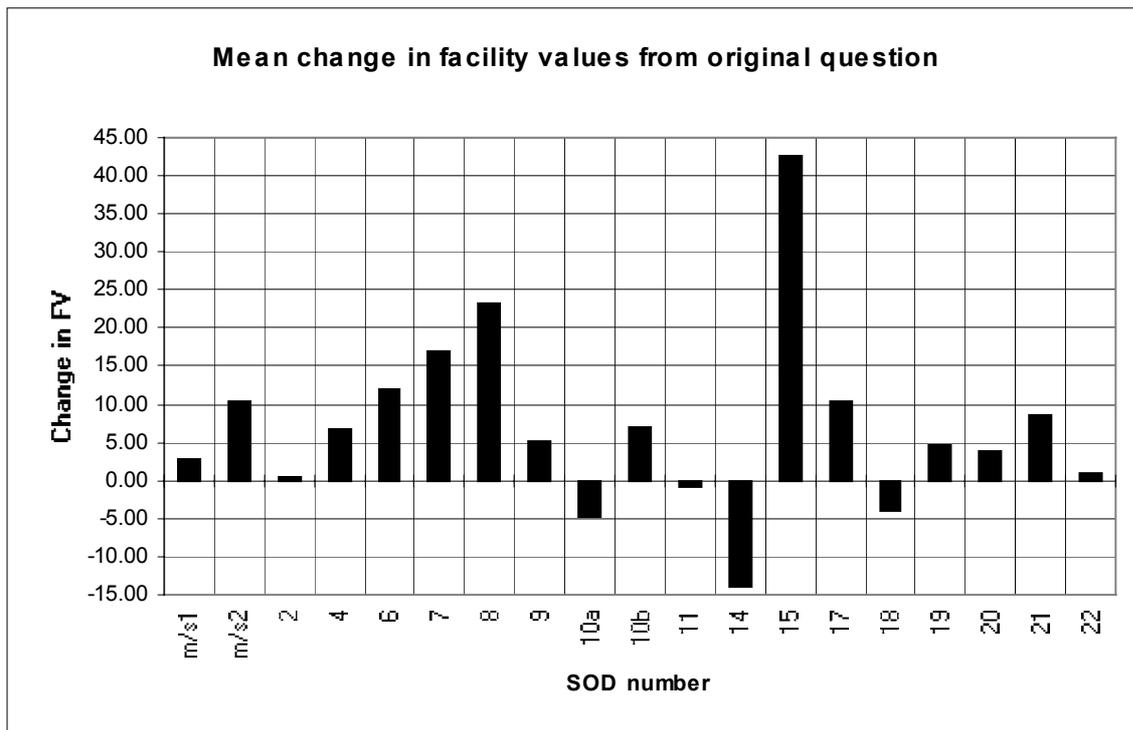
Graph to show the mean change in performance for each SOD manipulated Tier Y



Graph to show the mean change in performance for each SOD manipulated Tier Z



Graph to show the mean change in performance for each SOD manipulated all Tiers



It should be noted that these figures only indicate that there was a change in performance, little can be concluded from the magnitude of the change. However, for those SODs showing a large change the effect on performance is more likely to be the result of the manipulation rather than other factors such as the preparation of the students. For some SODs there were several questions that were changed and for others only one question, this table only gives a rough overall view of the changes that occurred. However, for each of the mean values calculated almost all the changes for any SOD were in the same direction, i.e. more difficult or easier. It is also important to note, that because of the imprecise nature of manipulating only one SOD at a time, some of the changes could be attributable to more than one change. It would be impossible to determine how much of any change was contributed to by any SOD simply by looking at the figures. Only qualitative analysis enables us to determine the cause of the change in performance.

Findings

Large changes in performance

The SODs that have shown large numerical changes are 6 (recall of strategy), 7 (recall of knowledge), 8 (number of steps), 14 (mathematical sequencing) and 15 (arithmetic errors). Although the magnitude of the change may simply be a symptom of the question chosen for manipulation, the size of the change indicates that this was most likely the result of the manipulation rather than other factors.

SOD 8 : Number of steps & SOD 15 : Arithmetic errors

SOD 8 and 15 are closely inter-linked: generally the more steps required in a calculation the more chance there is to make an arithmetic error. Figs. 8 and 9 show an original question and a manipulated version from Paper 4.

Fig. 8 : Paper 4 : Question 5 : Original question

5	Look at the sequence	2	8	18	32	
(a)	Write down the next term of the sequence.					[1]
(b)	Write down the 25th term of the sequence.					[1]
(c)	Write down the n th term of the sequence.					[2]

Fig. 9 : Paper 4 : Question 5 : Manipulated version

5 Look at the sequence

<i>Position</i>	<i>Term</i>
1	2
2	8
3	18
4	32
n	?

(a) Write down the next term of the sequence. [1]

(b) Write down the 10th term of the sequence. [1]

(c) Write down the n th term of the sequence. [2]

Part (b) has been changed to ask for the 10th rather than the 25th term of the sequence. This has reduced the potential for arithmetic error by reducing the number of steps that the student has to make to arrive at the answer when calculating the problem the ‘long way’ (which many did). This question also demonstrates the knock on effect of manipulations given that a table was added to assist with part (c) which may also have influenced the performance on part (b).

SOD 6 : Recall strategy & SOD 7 : Recall knowledge

Similarly the changes in performance from SODs 6 and 7 are also notable and linked. To help ‘recall of strategy’ the starting point for the calculation may have been given, for example starting a trial and improvement problem. Unsurprisingly, giving the student knowledge that they would otherwise have to recall, for example the formula for the area of a circle, makes the problem easier. In most of these cases this recall is a necessary part of the National Curriculum and therefore it would not be possible to give such formula in a live examination. However, the result is still interesting, because it can be seen that in most cases the students do not have problems with how to go about the calculation once they have the information. This may be particularly pertinent at the lower Tiers of entry. Geography and Classics examiners have confirmed that lower Tier candidates have problems with direct recall and consequently do not ask them to give examples. In both subject areas they have noticed that candidates find evaluating a given source much easier than direct recall. Consequently, in lower Tier Geography GCSE, a question will be phrased in such a way that candidates are told they ‘may use an example’. If a candidate gives a correct example this will be credited but they can still gain full marks without giving an example.

SOD 14 : Mathematical sequencing

A few of the manipulations created more difficult questions. This was particularly noticeable with the mathematical sequencing, where the order of the question was changed. This shows that the order in which a problem occurs is clearly important, which, in the case of Mathematics, is not surprising. It also demonstrates that the examples that were used were clearly in the most appropriate order before they were manipulated.

Negative change in performance

Other questions which increased in difficulty after manipulation were those with SODs 10a (context removed from data handling questions), 18 (spatial representation) and 11 (stated principle).

SOD 10a : Context (data handling)

Data handling questions that have no context are likely to be more difficult because they do not make sense. However, it could be seen for other types of questions the removal or change in context makes the questions easier.

SOD 18 : Spatial representation

It is more surprising that when the need for spatial representation was removed the question became more difficult. Adding a diagram of a four sided dice to a question that had previously not contained a diagram did not improve performance as might have been expected. This could have been because, despite the diagram, the concept of a four sided dice was still too unfamiliar or perhaps the diagram was not a good representation.

SOD 11 : Stated principle

Similarly, performance did not improve when the topic of the question was stated. Giving the topic did not give any information as to how to go about the question and this may explain the lack of improvement.

Appendix II (page 19) shows an example for each SOD that was manipulated and the associated change in Facility Value from the original version. Where the question appeared on more than one Tier the changes at each Tier are given.

Further investigation

It can be seen that, in most cases, the performance on any SOD is affected in the same way regardless of the Tier on which it appears. However, there are some exceptions to this, i.e. SODs m/s 2, 9, 17, 19 and 20. These changes will require further investigation to identify the questions manipulated and consider the reason for the differences in the direction of change.

Another area for consideration are the questions that appear across more than one Tier. From a quick glance it can be seen that the change in performance on the same question on different Tiers is not always comparable.

Conclusions

The results of the trial have shown that differences in performance can be influenced quite significantly by small variations in the questions. The insight into the changes these manipulations have made can help examiners to control difficulty of questions. Ensuring that questions are as difficult or as easy as the examiner intends will help to make them fairer for the candidates sitting the examination.

The next phase of the research will analyse the data from the trials in other subjects. Once the full analysis has taken place the results of all the subjects can be brought together and generalisations made where possible. Having done this, the next stage will be to create training materials for examiners which will enable them to use the information that we have gathered in their question writing process.

References

- Griffiths and McLone (1989) In *'Standards in advanced level mathematics Report of study 1. A study of the demands made by the two approaches to 'double mathematics'. An investigation conducted by the Standing Research Advisory Committee (SRAC) of the GCE Examining boards.'* UCLES.
- Hughes, S. and Fisher-Hoch, H. (1996) *What makes mathematics examination questions difficult?* Paper presented at BERA 1996.
- Pollitt, A., Hutchinson, C., Entwistle, N. and de Luca, C. (1985) *What makes examination questions difficult?* Scottish Academic Press.

APPENDIX I

SODs identified for each subject

Mathematics SODs

Source of Difficulty	Description
Command words	The words that tell the candidate what to do (e.g. 'explain', 'find', 'estimate', 'state', etc.) were not always the same across questions.
Context	The scenario in which the question was set could inhibit the development of a mental model of the question. If the context was inaccessible then the Mathematics was often also inaccessible.
Stated principle	If the mathematical topic or concept was not given then candidate had to deduce which topic the question related to.
Combination of topics	Difficulty arose in questions which involved more than one mathematical topic.
Isolated skill or knowledge	The area of mathematical knowledge or skills required was not well practised by the candidate because it did not overlap with other syllabus areas.
Mathematical language	Recognising mathematical terms.
Mathematics v. everyday language	Mathematical and everyday language could have different meanings.
Mathematical sequencing	The sequence of the sub-parts of the question did not always follow appropriately.
Recall strategy	This was exacerbated when there was a need to recall a strategy that was not given. If the strategy was not recalled then devising a strategy could be more demanding.
Alternative strategies	Alternative strategies to those anticipated by the examiner could require more steps. This required more of working memory capacity. This had implications for the allocation of marks, where a mark scheme had not anticipated the use of alternative strategies.
Abstraction required	Abstract thought was required.
Spatial representation required	Spatial skills were required to build a mental model of the question.
Paper layout	Physical organisation of the question ordering and or numbering could support or hinder candidates.
Ambiguous resources	Unclear resources (diagram , graph , table etc.) affected performance.
Irrelevant information	Information appears in question that was not required may have distracted from relevant information.
Number of steps	A large number of steps over-loaded working memory and information was likely to be lost.
Arithmetic errors	Some questions had more opportunity for making arithmetic errors than others.

Geography SODs

SOD/E	Comments
<i>Accuracy</i>	inaccuracy of measurements
<i>Choice</i>	differences in difficulty between choice of questions
<i>Command words</i>	single words e.g. describe, explain etc.
<i>Distractors</i>	too much information can distract candidates
<i>Emphasis</i>	use of bolds, capitals, italics
<i>Example</i>	asking directly for an example
<i>Expectation</i>	answering to the question expected, not the actual question
<i>Geographical knowledge</i>	answer unknown, regardless of the question
<i>Incompatibility</i>	mark scheme and question are not compatible
<i>Information</i>	insufficient information given
<i>Instructions</i>	Instructional phrases e.g. 'Give two reasons...'
<i>Mark allocation</i>	breakdown of mark allocation given on the paper
<i>Paper layout</i>	space allocated for responding
<i>Response prompts</i>	words or phrases that appear at the beginning of the answer space
<i>Sequence</i>	order of question bits within the question
<i>Structure</i>	open endedness of the question
<i>Topic related knowledge</i>	other knowledge about the same topic interferes with the response
<i>Words</i>	meaning of words unknown or unclear

Science SODs

Source of Difficulty	Description
Mis match between intention and question	The question that is asked is not the one the examiners want them to answer.
Level of explanation/detail required	The question does not make it clear OR candidates misunderstand either how much detail is needed or what level of explanation is required.
Alternative answer	There are cases where alternative answers are (i) correct (almost); (ii) possible; (iii) plausible; (iv) suggested by the context.
Misunderstanding word/grammar/emphasis	These are cases where the errors suggest the candidate has not understood what the question was asking for.
Context related	These include cases where (i) context cues different answer/more familiar strategies; (ii) context unfamiliar; (iii) context contains distractor.
Interference for topic-related knowledge	Candidates influenced by other knowledge on the same or a related topic.
Everyday knowledge	Everyday knowledge is necessary or interferes.
Specific knowledge/recall required	Knowledge to be recalled is very specific.

English SODs

Description	SOD	SOE*
In reading the question rubric and understanding the task	There is question choice	There is no question choice
	Question rubric is unclear	Question rubric is clear
	No highlighting emphasis	Highlighting emphasis
	Technical language	No technical language
In finding the correct piece(s) of text, from which an answer might be derived;	Text reference given is general e.g. paragraph	Text reference is specific e.g. word
	Text reference given is not where answer actually is.	Text reference leads exactly to relevant piece of text.
	Relevant text is in different place to last question	Previous question uses the same text reference
	Text reference needed	Need for text reference removed by giving all necessary bits of text in the question.
In understanding the meaning of the identified text piece(s) at the level either of decoding or interpreting;	Word has more than one meaning, depending on context	Word has only one meaning
	Idiom was taken literally	No idiom involved
	Text and questions are in L2	Text and questions are in L1
	Rare or difficult vocabulary	Vocabulary is familiar or easy
	Dealing with abstract concepts	Dealing with concrete descriptions
In the composition of an adequate written response.	Context (audience) for the response is not given.	Purpose of writing (audience) is clear
	'Own words' required	Candidate can lift some words from text without translating into their own words
	Text needs manipulation for answer (grammatical, syntactical)	Candidate can lift chunks of text without changing it in any way.
	Production of difficult vocab	Recognition of difficulty vocab
	Quantity of response is larger (number of words or reasons)	Quantity of response is smaller
	Comparing required	Description required
Marking	Mark scheme for summary is linear	Mark scheme for summary is hierarchical
	Outcome space is clearly defined	Outcome space is uncontrolled
	Mark scheme is rigid	Marks scheme is flexible (allows a variety of answers)
	Mark allocation does not reflect the expected number of parts to the answer	It does
SODs occurring throughout the process	Opening up the question	Closed questions

*SOE : Source of Easiness

French SODs

SOD	SOE	Notes
Vocab is difficult in the question in the text	Vocab is easy in the question in the text	
Procedural knowledge is required (i.e. not language knowledge)	Language knowledge required only	Is this valid? What are the aims of the syllabus?
Need to read the text - information is not available by guessing or common sense	No need to read the text - answers can be guessed correctly with common sense	Are the pieces of text adjacent? how close are they?
Question text reference is semantic - a synonym is used to match piece of text - but no structural cues	Question reference is structural - exact match physical of work or number	This is in a continuum - the level of the text reference.
The answer is in two parts of the text and you need both to get the answer.		
The answer is in two parts of the text and you only need one part to get the answer		This will therefore give two attempts at answering - so is it a SOE or confusing?
A text composed of 6 questions with 6 possible answers did not require that each answer was used once.		Try more questions or fewer people, don't match number of questions and no. of possible answers.
A multiple choice with 6 options, but the 6 options were not discrete enough - there was some overlap in content.		Are the options discrete? What did people do? Did they take answers from the under-used parts of the text?
The questions focus on one part of the text and there are not questions on another part of the text.		
Question required inference not just straight comprehension of the text.		A continuum of the degree of difficulty/inference.
Need to flit about around the text	Questions are in the same order as the information in the text	Is this valid? Try sequencing questions to match text and not sequencing.
MS asked for 2 points, but the points are closely related that they may be seen as the same point.		i.e. there is no inference expected, but the two points are linked.
Higher mark allocation was given where questions were assumed to be more demanding, but it seems that they were not any more demanding. (maybe this reflects the objectives of the test?)	More demanding questions were not given higher mark allocation	Weighting the type of questions.
A section of questions do not work as separate questions, but as one question with a high mark allocation (this is a text effect).		Yes, but it is better to have a number of texts to spread the risk of this throughout test.

French SODs (cont.)

Question is about the general theme of the text therefore require understanding of the gist of the text		
MS rewards only one of a number of possible answers.		
Synonym in question (for word in text) may not have quite the same meaning.		
Need to transform text to compose an answer	Can lift answer straight from the text	
True/false questions discourage reading of text.		But it may have been the nature of the question that changed the nature of how it was read. Did they read the questions first, or after the text?

APPENDIX II

Examples of SOD manipulations

SOD	Frequency of manipulation	Paper	Question Number	Question Bit	Version
m/s1	5	2	2		e
m/s2	4	2	16	b	e
1	not manipulated				
2	6	2	13	a&b	c
3	not manipulated				
4	8	3	5	a	c
5	not manipulated				
6	2	4	6	b	b
7	9	2	17		d
8	1	3	4	b	b
9	3	4	13	c	d
10a	8	4	13	a,b&c	c
10b	29	3	15	c	b
11	3	4	12	a,b&c	d
12	not manipulated				
13	not manipulated				
14	3	3	13	a&b	b
15	1	4	5	b	b
16	not manipulated				
17	4	2	17		c
18	3	1	20	a,b&c	d
19	5	6	10	a,b&c	d
20	4	5	14	aii	c
21	16	1	19	a&b	b
22	1	4	1	b	d

The next pages show these examples, giving the original question, the manipulation made and the associated changes in Facility Value (F.V.). Where the questions appear on more than one Tier the change in F.V. is shown separately.