Paper presented at IAEA Conference, Slovenia, May 1999.

Curriculum Demands and Question Difficulty

Ayesha Ahmed and Alastair Pollitt

Introduction

In order to improve the quality of examination questions and ensure that they make the intended demands on students, we need to understand the psychological processes involved in answering questions. We have developed a psychological model of the question answering process and a scale on which demands can be measured. We have also carried out empirical work in five subjects in order to identify factors that affect question difficulty. This paper is an attempt to integrate the model, the scale of demands and the work on question difficulty, in order to give a fuller and more coherent picture of demands and difficulties in examination questions.

Model of Question Answering

The model of the psychological processes involved in answering examination questions consists of six stages. The first stage, Stage Zero refers to the student's understanding of the subject knowledge, that is what the student has learned before the examination. For example, the student's understanding of Geography, or of History, or of more specific topic areas within subjects.

Stage One is when the student forms a mental model of the question, by reading the question and forming a representation of it. This representation could be an understanding of aspects of the question and it could also consist of a misunderstanding of aspects of it. Stage Two is the process of searching the subject knowledge for relevant concepts to use to answer the question.

Stage Three consists of an attempt to match the representation of the question with the aspects of subject knowledge that have been selected during the search process. Stages Two and Three, searching and matching, may be simultaneous in some cases. In Stage Four the student generates an answer and in Stage Five the student writes an answer. Again in some cases these last two stages are simultaneous. All the stages happen in quick succession, often overlapping one another and sometimes subsuming one another.

Sources of Difficulty and Easiness

This model has been developed after empirical work on the Sources of Difficulty and Sources of Easiness (SODs and SOEs) in examination questions based originally on work by Pollitt et. al. (1985). The evidence from these studies tells us which aspects of questions affect difficulty, and these SODs and SOEs can have an effect on difficulty at various stages in the process of question answering outlined above.

We identified Sources of Difficulty and Sources of Easiness through empirical work in five subjects. Scripts from examinations in Mathematics, Geography, Science, English Language and French taken at age 16 were analysed statistically in order to identify the most difficult questions. The students' answers to these questions were then analysed with the aim of discovering the most common errors made when answering these questions. The common errors were used to help us to identify what was causing students to lose marks on the questions, and from these errors we were able to hypothesise that there were certain Sources of Difficulty and Easiness in the questions.

In order to discover whether the SODs and SOEs did have an effect on performance, we manipulated various aspects of the questions and trialled these manipulations in schools. The manipulations were based on the hypothesised SODs and SOEs in the original questions. SODs and SOEs were systematically added or removed from the new trial questions so that we could measure their effect on the performance of students. A list of SODs and SOEs that affected question difficulty was produced for each subject, and from this we put together a list of general SODs and SOEs that applied to all five subjects and that can be applied to almost all subjects.

SODs and SOEs can be either valid or invalid. A valid SOD or SOE is one that the question setter intended to put in the question. An invalid SOD or SOE is one not intended by the question setter. It is likely that the question setter will not be aware that this SOD or SOE exists in the question, and it will therefore pose a threat to construct validity. The students answering the question may answer in a way unanticipated by the setter and therefore unrepresented in the mark scheme. Students may misunderstand the question and therefore not be able to show the marker what they know about the topic. The cognitive processes that the setter intends the question to provoke may never occur in students' minds.

The SODs and SOEs have an effect at different stages in the question answering process. Some can affect the cognitive processes in students' minds at more than one stage. The SODs and SOEs are listed below, grouped into the stages in which they might have an effect:

Stage 0 – Knowledge of the subject
Concept difficulty - abstractness or unfamiliarity of concepts

Stage 1 – Understanding the question
Distractors - in wording of question
Context
Highlighting - of words or phrases
Density of presentation
Technical terms
Everyday language
Inferences necessary
Command words

Stage 2 – Searching subject knowledge
Sequence of questions
Combination of topics
Resources

Stage 3 – Matching representation of question and subject knowledge
Context

Stage 4 – Generating an answer
Density of presentation
Inferences
Mark allocation
Response prompts - for content of answer
Command Words

Stage 5 – Writing an answer
Response prompts - for organisation and structure of answer
Paper layout
Own words - transform text into your own words
Whole resource processing – writing a summary of a text

Scale of Cognitive Demands

In parallel to discovering how and where SODs and SOEs affect performance in examinations, we also investigated the issue of what makes questions demanding. We had identified various factors that caused difficulty, but found that questions that are difficult can make demands on students in different ways. In an attempt to describe the different ways that demands are made on students we developed a scale of cognitive demands which had four dimensions: Complexity, Abstraction, Resources and Strategy.

The SODs and SOEs affect students at various stages in the question answering process, and the level of demands on each of the four dimensions in the scale reflect how many SODs and SOEs occur in a question and how likely they are to affect the process of question answering.

The scale of cognitive demands was developed from a scale devised by Edwards and Dall'Alba (1981). The original scale used by Edwards and Dall'Alba was designed for rating the demands in science lessons. We adapted this scale for use in rating the demands of examination questions in all subjects (see Pollitt, 1998). The four dimensions in our final scale: Complexity, Abstraction, Resources and Strategy were used successfully by examiners in Geography, History and Chemistry, to rate the demands of questions. These ratings are described in detail in Hughes et. al. (1998).

The complexity of a question is best described by the number of operations that have to be carried out, or the number of ideas that have to be brought together, and also the nature of the relationships or links between them. A question that has a low complexity,

would involve simple operations or ideas and no need to link them together. There would also be no real demand for comprehension of concepts – just of the question itself. A question with a high level of complexity would involve combining operations and ideas, linking them and evaluating them, with a need for the comprehension of technical concepts. A complex question can be an easy question, but it cannot be a simple question.

The next dimension refers to the resources provided with the question such as a text, diagram, table, picture, graph or photograph. It also refers to the internal resources that students bring to a question, that is their mental representation of their subject knowledge. Questions with a low demand in terms of resources will direct students as to how to use a printed resource, and which parts of it to use. The resources will contain all the information students need to answer the question so they will not have to make inferences from it. It will also contain only relevant information, so that nothing has to be filtered out or inhibited by the student. A question with a high demand in terms of resources will require students to generate the information needed to answer the question from their internal resources, or to select, interpret and make inferences from any resources provided.

Abstraction refers to the extent to which the student has to deal with ideas rather than with concrete objects or events to answer the question. Students' understanding of theoretical issues is tested, rather than their knowledge of specific examples. A question with low demands on this scale will involve working with concrete concepts. It may focus for example on events in a History question or in an English Literature question, or on specific places in a Geography question, or specific experiments in a Science question. A question with high demands on this scale will involve highly abstract concepts and will also require a grasp of the technical terms referring to these more abstract concepts.

The strategy dimension refers to how students go about answering the question. It refers to how students devise a method for answering, and how they maintain their answering strategy, monitoring it as they go along. They may have to devise a new strategy to answer the question or they may select a strategy they have used before. The strategy is applied to the process of tackling the question and to the process of communicating an answer. Some examiners found it useful to separate strategy into two dimensions: task strategy and response strategy. The task strategy is the one which the student applies to the process of tackling the question, and the response strategy is applied to organising and communicating an answer.

A question with low demands in terms of task strategy would provide a strategy so that students do not need to devise or monitor one themselves. High demands for a task strategy would arise when students are not given any support with how to tackle the question. This is common in unstructured questions in which no prompts are given and students have to form their own strategy for how to approach an open-ended task. Low demands for a response strategy would occur when students do not have to organise their own response, but are given prompts on how they should structure an answer. High demands for a response strategy would occur when students have to select appropriate

information for their response and organise this information themselves, and again this is more likely to occur in a more open-ended question.

Demands and Difficulty

We can now consider how these demands relate to the SODs and SOEs. The more demanding questions may have a larger number of SODs that have a significant effect on performance.

Complexity

The first dimension, complexity, can be thought of in terms of the hurdles students have to get over in order to produce an answer to a question. The level of complexity would correspond to the number of hurdles and also the height of the hurdles and such issues as whether students have to get over one first in order to attempt the next.

Complexity affects the process of reading and understanding or misunderstanding the question (Stage One). Students are more likely to misunderstand a complex question, especially if complex language is used. A complex question will cause more concepts to be activated in students' minds and there is therefore a greater chance that students will arrive at incorrect concepts when searching their subject knowledge and matching this with the concepts elicited by the question. This will result in more mismatches between the concepts in the question and the student's subject knowledge. As well as the complexity of the question, the actual task to be carried out can be complex, requiring recall and evaluation of knowledge rather than just recognition processes. The command words used in the question, which can be a SOD, influence the process of understanding the question and generating an answer. Command words such as 'Explain' and 'Evaluate' will result in questions with a higher level of complexity than command words such as 'Describe' or 'State'.

Complexity also affects the response process, by bringing together Stages Four and Five of the model of question answering. If the response students have to give is complex then these two stages become one because students have to generate an answer and think about how they are going to communicate it at the same time. The process of generating an answer is more conscious when a question has a high level of complexity, and the process of writing the answer simply falls out of this. In this case, communicating the answer is part of the process of generating it.

However, if the answer students have to give involves a simple procedure and the concept that they need to use emerges automatically from the process of matching their knowledge with information in the question, then Stages Four and Five are separated, and there is more of interest happening in the writing process (Stage Five). The cognitive effort goes into the process of organising your answer and deciding how you are going to communicate it, rather than into generating an answer.

Different psychological processes therefore occur in students' minds depending on the level of complexity of a question. If the answer is complex, conscious processing occurs when generating the answer, whereas the process of communicating the answer is an implicit part of this. However, if the answer is simple, generating it is a more implicit process, but communicating the answer involves conscious and explicit processes.

Abstraction

The process of understanding the question and answering it can be complex, but also the actual concepts necessary to answer the question can be complex and this is where abstraction comes in. When the concepts themselves are complex they are often highly abstract, but abstract concepts are more than just complex concepts, they are also complex in the way that they are represented in the mind. This is related to Stage Zero of the model of question answering, when students are learning and forming a representation of the concepts involved in the subject. The SOD that we call Concept Difficulty occurs in questions that are highly demanding in terms of the abstraction of concepts.

When students form a representation of a particular subject area it will involve representing some abstract concepts and some more concrete concepts. The abstract concepts may be represented in a different way from the more concrete ones. Abstract concepts are linked to representations of events, that is particular concrete experiences which are represented using language, but they do not refer directly to concrete objects or events. Because they are represented using language, these concepts have to be consciously learned and therefore they often have to be taught (Kintsch, 1998).

Concepts that are abstract can be concepts that are distant in time and place and from experience. For example, in a Geography examination in which questions were rated on the scale of cognitive demands, the highly abstract questions were not about the here and now but about geographical events in the past, or how to predict them in the future, and about places distant from the students' experience. If a question requires a generalisation so that students have to write about something theoretical when given a specific example, then the question will be highly demanding. Deriving a concept or theory from an example is more difficult than deriving an example from a general or theoretical statement.

Students need to learn the meaning of abstract concepts within each subject area, but these are much more difficult to understand than concrete concepts as they have no concrete referent. In some cases a verbal definition can be given for students to learn and in some cases a concrete example can be given. However, in the most difficult cases there will not be an example that will be known and understood by everyone, so there cannot be an agreed definition or instantiation of that concept. In these cases each individual builds up their own idea of what the concept means, and this is hard to verbalise.

Our abstract concepts are linked to our own particular experiences, so one person's idea of the meaning of a concept will differ from another person's idea of the meaning of the

same concept. Each person links their representation of that concept with their own experiences, which are different from other people's experiences. Abstract concepts for which there is no shared definition will therefore make a question highly demanding. If students are required to explain why someone is 'brave' or 'wise' or if something is 'good' or 'efficient' then they cannot simply make a statement, but have to form an argument and explain what they mean to the marker who has his or her own representation of what these abstract concepts mean. As well as convincing the marker, the student has to convince the question setter who has written the mark scheme.

Another result of the way abstract concepts are represented is that their representations will be linked to many different events, so they will be highly associative. Because of this, when an abstract concept appears in a question it is more likely to cause the activation of many other concepts in students' minds. Students are therefore more likely to use different concepts from those intended by the examiner when answering the question. As well as having more links to other concepts, abstract concepts will have weaker and more distributed links than concrete concepts. The danger of a question causing the activation of incorrect or inappropriate concepts is therefore greater when the question contains abstract concepts. This also means that abstract concepts are more difficult to access themselves, as they are only weakly linked to other representations. Similarly, if a question refers to an isolated topic that is not well connected to other concepts then it will be more demanding as students cannot arrive at the answer so easily via other concepts. Concepts that are infrequently used are not very well linked with other concepts, and neither are concepts that are unfamiliar because they are distant from experience. Abstract concepts therefore differ from concrete concepts in the complex way that they are represented in the mind.

All of these factors combine to make a question involving abstract concepts a highly demanding question. Once the correct concepts are arrived at in a question, students then have to translate these ideas into a written answer (Stage Five of the model). This stage is also more demanding when abstract concepts are involved as students have to make an argument and explain what they mean. Abstract concepts are also often expressed in difficult vocabulary or technical terminology. These words can act as invalid SODs if the setter is intending to assess students' understanding of the concepts and not of the terminology. A further SOD can occur when a word means one thing in everyday language and quite another in the context of the topic being assessed. The word will trigger students' schemas for its meaning in everyday language and this can cause misunderstanding, or distract them from answering the question. To use the word correctly students must remember to make the technical meaning override the everyday default meaning.

Resources

The dimension of resources can also affect the way in which Sources of Difficulty act during the question answering process. In order to be able to use resources it is necessary to know what resources are available. When the resources are printed within the examination paper then this task is easy. However, students also have to use their own

resources to answer a question. They have internal resources consisting of their knowledge of the subject. They have memories of particular things they have learned about the subject and memories of the context in which they learned things: the teacher, the room, the emotions, the common sense knowledge that relates to the topic. The internal resources that students could potentially use to answer a question are unbounded. Their knowledge is linked to so much other knowledge and to many emotions and situations and aspects of the environment. The real task is for students to select the relevant resources. They have to select which information out of their ever-expanding network of knowledge is appropriate to use for answering this particular question. Knowing how to use internal resources is therefore an important aspect of answering an examination question.

The difficulty with selecting the correct internal resources to use when answering a question is that there will be many irrelevant aspects of knowledge that are activated by a question. This is an even greater problem when there is information in the question that may lead to students selecting the wrong resources to answer it. For example the question may be in a context that is particularly distracting, and leads students to use the wrong part of their knowledge. If aspects of the context, particular words or phrases or even a picture, are particularly salient they are more likely to cause activation of representations of irrelevant knowledge. The context and distractors in the question are both Sources of Difficulty identified from our empirical work, and they are invalid SODs if they are preventing the question from assessing what the examiner intended by causing students to select the wrong information to use to answer the question.

In normal circumstances such problems are solved or avoided by monitoring, that is by attending to the task. When attention is focused on the task and its outcomes, irrelevant information is inhibited successfully and students can answer using only the relevant knowledge. However, in an examination situation students are under stress. They may be anxious about the examination, and their ability to inhibit irrelevant information may be affected by this.

Fox (1994) has found an attentional bias in highly anxious people, which results in them allocating attentional resources to particularly distracting information and failing to inhibit it successfully. This can then disrupt the processing of subsequent information. Students may be highly anxious in an examination situation and this will affect their processing of questions and could result in errors. A further problem occurs when particularly emotional words are used in a question. Students are likely to allocate too much of their attentional resources to these words. The question may be mis-read or mis-interpreted because the student's attention is focused on the salient words. This can also be a problem in questions where there is a context that is distracting. In this case, if words in the context are highly salient, and the students are anxious about the examination, defective inhibition of the distractors could occur and cause them to make errors in the reading and interpretation of the question.

Focusing attention is the key to correct interpretation of an examination question. Attention is easily distracted by salient words or contexts in a stressful examination

situation, and this causes activation of incorrect representations (See Evans, 1989 for a discussion of the role of selective attention in causing errors).The SOD called Highlighting can modify the students' attentional processes by making a particular word or phrase more salient. This can help students to focus their attention on what the question is asking. Another SOD, Density of Presentation, can affect the students' attentional processes, making it more difficult to select and focus on relevant information in the question.

The issue of knowing your own internal resources is also crucial: knowledge of what you know, or metaknowledge, is necessary so that you can then use this knowledge in the appropriate way. Examination questions demand that students know what they know, and know how to use this knowledge and when to use this knowledge. They also require that students know their own strengths and weaknesses, especially when there is question choice.

The student's internal resources can be thought of as an extension of the resources on the page. Internal resources have to be searched as described in Stage Two of the model of question answering, and resources printed on the examination paper also have to be searched in order to identify relevant information.

Strategy

The fourth dimension, strategy, may be thought of as including one strategy for tackling the question and one for communicating the response. These correspond to Stages Four and Five of the model of question answering. The task and response strategies are therefore brought together to varying degrees depending on the complexity of the question.

Strategies are affected by the existence of SODs and SOEs in the question. Response Prompts for the content of an answer can influence the task strategy that students use by provoking them to use certain concepts in their answers. Similarly, Response Prompts for the organisation and structure of an answer can influence students' response strategies. Questions with Response Prompts tend to be less demanding than more open-ended questions as they suggest strategies to the students. However, for the more able students these questions can be restrictive. Other SODs that can influence response strategies are Paper Layout and Mark Allocation. The amount and nature of the space given to students for answering a question and the number of marks available can affect the amount students write and the level of answer they write.

The strategies students use are also affected by the command words given in the question. When students are asked to give an explanation their strategy will be different from when they are asked to give a description. A further problem students face is how to interpret command words such as 'explain' in the different subjects they are studying. 'Explain' in Geography means something different from 'explain' in History or in Mathematics. Students have to interpret the command word in the context of the subject and even

sometimes of the topic within a subject. Once they have recognised the kind of explanation they have to give, they then apply the appropriate strategy.

Conclusions

The Scale of Cognitive Demands along with the Model of Question Answering can give us a clearer understanding of the cognitive processes occurring in students' minds when they are answering examination questions. The Sources of Difficulty and Sources of Easiness are placed in questions either intentionally or unintentionally by examiners, and these have an effect on the students' cognitive processing at various stages in the Model and to differing degrees. Questions rated as having high demands in terms of the four dimensions on the Scale tend to have more SODs and fewer SOEs occurring at the various stages.

An understanding of these cognitive processes can help question setters to write questions that will make the demands they intend. This allows us more control over the Sources of Difficulty and Easiness in questions and helps to ensure the validity of the demands that the questions make on students.

References

Dall'Alba, G. and Edwards, J. (1981*)* The scale of cognitive demand: An instrument for analysing cognitive demand in secondary science. *Educational Research and Development Unit. Royal Melbourne Institute of Technology, Australia.*

Evans, J St.B.T. (1989) *Bias in Human Reasoning: Causes and Consequences*. LEA, Hove, UK.

Fox, E. (1994) Attentional Bias in Anxiety: A Defective Inhibition Hypothesis. *Cognition and Emotion*, 8 (2), 165-195.

Hughes, S., Pollitt, A., & Ahmed, A. (1998) The Development of a Tool for Gauging the Demands of GCSE and A level Examination Questions. *Paper presented at BERA, Belfast.*

Kintsch, W. (1998) *Comprehension. A Paradigm for Cognition*. Cambridge University Press.

Pollitt, A., Entwistle, N., Hutchinson, C. and De Luca, C. (1985) *What Makes Exam Questions Difficult?* Scottish Academic Press, Edinburgh.

Pollitt (1998) Maintaining Standards in Changing Times. *Paper presented at 24th Annual IAEA Conference, Barbados.*