

## **The development of a tool for gauging the demands of GCSE and A Level exam questions.**

Sarah Hughes, Alastair Pollitt and Ayesha Ahmed.

University of Cambridge Local Examinations Syndicate,  
Hills Road, Cambridge, CB1 2EU. Phone 01223 553845

### **ABSTRACT**

Two techniques from the psychological literature were used to identify demands in exam questions: Edwards' scale of cognitive demand (1981) and Kelly's Repertory Grid technique (1955). These two techniques were used to develop a tool for identifying and gauging the demands made in GCSE and A Level History, Chemistry and Geography questions. This paper reports on three phases of the development: Phase 1, the adaptation of Edwards' scale to assessment tasks in a number of subjects; Phase 2, subject specialists' elaboration of the scale and; Phase 3, the integration of examiner's perceptions of demands in exam questions.

### **INTRODUCTION**

Demands are requests to perform made of candidates by examiners. There can be many or few demands in a question, and these demands may be complex or simple. Skills, knowledge, understanding and the ability to apply these are generally seen by examiners as the demands of exam questions. Given that there are many different sorts of demand in examination questions this paper refers mostly to *demands* rather than *demand*.

Sometimes demands are described euphemistically as 'opportunities', where credit will be given to those who can meet them. Sometimes they are explicit in the question or task, but sometimes they reside in the mark scheme and it is the job of the teacher to ensure that pupils are fully aware of them. In recent years the use of grade and level descriptors in the more literary subjects has increased the awareness amongst pupils and teachers of the implicit demands residing in the mark scheme, and it has become clear that these demands in fact constitute a significant part of the syllabus.

In a study of A level maths (SRAC 1990) it was noted that skilled judges (ie examiners) are able to recognise demand and generally agree with each other in estimating the overall level of demand in questions. However, although they could recognise 'demand' they were much less good at explaining it; they could say, and agree, that a task was difficult but could not analyse it into the cognitive elements and processes that were the source of that difficulty. This should not be seen as a criticism of the judges, since they were mathematicians not psychologists, but if we are to arrive at a proper explanation of the demands and difficulties of exam questions, and so to achieve control of this most central element of examining, we need to bring together the expertise of both the subject specialist and the psychologist. In this paper we report the use of two techniques from the psychological literature - Edwards and Dall'Alba's (1981) scale of

cognitive demand and Kelly's (1955) repertory grid interview - to help the development of an instrument for identifying and quantifying the demands in A level and GCSE questions in History, Chemistry and Geography.

### **EDWARDS' SCALE OF COGNITIVE DEMAND**

Edwards and Dall'Alba (1981) developed and implemented a 'Scale of Cognitive Demand'. The scale was developed as an instrument for analysing secondary science lessons, materials and evaluation programs in Australia. It was intended to quantify the demands placed on the cognitive abilities of students. The conceptualisation of demand was derived from a range of learning and thinking theories, including Bloom (1956), Burner et al (1966), de Bono (1976), Gagne (1970), Taba 1962, 1967, Ausubel (1978) and the work of Piaget as interpreted by Novak (1977). The scale considered demand to have four dimensions: Complexity, Openness, Implicitness and Level of Abstraction. Six levels of demand were defined within each dimension, by a list of phrases and command words that were typically used in science textbooks and examinations, or that could be used to describe the processes students were required to carry out. The original scale is shown in figure 1.

**Figure 1 The Scale of Cognitive Demand Dall’Alba and Edwards (1981)**

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
<b>Complexity</b> The nature of the sequence of operations which constitutes a task, that is, the nature of each component operation and the links between operations.	Simple operations	Require a basic comprehension	Understanding, application or low level analysis	↔	Analysis and/or synthesis	Synthesis or evaluation
<b>Openness</b> The degree to which a task relies on the generation of ideas.	No generation of new ideas	↔	Limited generation of ideas	Generation of ideas from a given data base	Generation of ideas which are original for the student	Highly generative
<b>Implicitness</b> The extent to which the learner is required to go beyond the data available to the senses.	Data are readily available to the senses	Data to be operated on are given	A large part of the data is given but requires generation of the final outcome	↔	Data are not available in a readily useable form - must be transformed	Require a view of the entity in question as part of a more extensive whole
<b>Level of Abstraction</b> The extent to which a task deals with ideas rather than concrete objects or phenomena.	Deals with concrete objects or data stored in the memory	Predominantly deals with concrete objects or images	↔	Corresponds to concrete-abstract transition	Abstract	Highly abstract
Key words	Recall or memorise specifics Assimilate information Simple measurement Observe Simple comparison Simple recording Follow a simple set of instructions Observe with discrimination	Recall or memorise procedures, processes, rules of principles Simple classification Demonstrate use of equipment Understand a direct, physical model Translate Summarise	Relate Develop an operational definition or simple concept Simple application Simple extrapolation Compare on stated criteria Identify discriminating characteristics	Internalise a concept Apply a rule of principle Classify Simple hypothesising Complete an experimental design Understand a model of abstraction	Construct a model or other representation Generate relevant criteria Extrapolate Generalise Hypothesise Isolate variables Design an experiment or piece of equipment Isolate inferences or assumptions Integrate	Develop or explain a concept of an abstraction Link a model with reality Assess the impact Evaluate

Edwards' original scale was modified to enable the scale to be applied to subjects other than science. The scale underwent a number of changes. These occurred in three phases. Phase one was the adaptation of Edwards' scale to enable it to be applied more readily to subjects other than science. Phase two involved some elaboration of the three subject scales by examiners who had used the scales. Phase three was the integration of data from Repertory Grid interviews with examiners. These three phases are described in this paper:

### **PHASE 1 - THE ADAPTATION OF EDWARDS' SCALE OF COGNITIVE DEMAND**

Phase 1 involved three adaptations of the scale:

First, the two dimensions Openness and Implicitness were merged. Openness refers to 'the degree to which a task relies on the generation of ideas' (Edwards and Dall'Alba 1981 p2). Implicitness refers to 'the extent to which the learner is required to go beyond data available to the senses' (Edwards and Dall'Alba 1981 p3). Both of the dimensions pertain to the use of resources - whether those resources are given (the data referred to in Implicitness) or internal (the knowledge and ideas referred to in Openness). In some subjects the distinction between given and internal resources is less defined. In history for example, document analysis tasks common in both A Level and GCSE involve the use of both given information (historical documents) as well as candidates' knowledge and understanding. The new category which incorporates these two things is called 'Resources'. It relates to the information given and how much candidates have to generate their own information, as well as what they do with that information.

Second, a dimension relating to the strategy students use to answer questions was added. There were no dimensions in the original scale which referred to (i) devising and maintaining a strategy for answering the question and (ii) communicating an answer, however strategic demands have been flagged as important to learners by, for example, Ausubel (1978) and Rigney (1978). Ausubel described how the provision of 'advanced organisers' which give a structure to masses of information and provide a central framework upon which learners can organise and elaborate their knowledge. Rigney (1978) described learning strategies as cognitive strategies which facilitate learning.

"Cognitive strategy will be used to signify operations and procedure that the student may use to acquire, retain and retrieve different kind of knowledge and performance. These operations and procedures may be cognitive information processing, as in mental imagery, or may be cognitively controlled, as in skimming through a text book to identify major points. Cognitive strategies involve representational capabilities of the student (reading, imagery, speech, writing and drawing), selectional capabilities (attention and intention) and self directional capabilities (self programming and self monitoring)."

Rigney 1978 p165.

Glaser and Bassok (1989) described expert learners as those who are able to select relevant information, select appropriate strategies for tackling the task and monitor and regulate their cognitive processes.

The demands of organisation and monitoring a strategy are the demands made by exam questions and should be incorporated in a tool which aims to measure demands. In GCSE and A Level questions the process of answering is important, and often carries marks (i.e. is one of the assessment objectives). For example, a GCSE Chemistry syllabus states that 15% of marks should be given for ‘evaluation and communication’.

The dimension ‘Strategy’ has been added to the scale, it refers to the demands of selecting or devising a method to use to approach the task, monitoring of the progress of the plan, and the organisation of the answer.

Third, we found it necessary also to change from six defined levels for each dimension to a 1-5 continuum with only levels 2 and 4 described verbally. The language used in the original scale was science specific; by keeping the dimensions less stringently defined examiners were more readily able to apply the scale to their subjects. Levels 1, 3 and 5 are not defined as this stage, this allowed the examiners to use their professional judgement to apply the scale.

The revised scale, known as the CRAS scale is shown in figure 2:

Figure 2 The ‘CRAS’ Scale of Demands

	1	2	3	4	5
<b>Complexity</b> The complexity of each component operation or idea and the links between them.	←	Simple operations (i.e. ideas/steps) No comprehension, expect that required for natural language No links between operations	↔	Synthesis or evaluation of operations  Requires technical comprehension Makes links between operations	→
<b>Resources</b> The use of data and information.	←	All and only the data/information needed is given	↔	Student must generate the necessary data/information.	→
<b>Abstractness</b> The extent to which the student deals with ideas rather than concrete objects or phenomena.	←	Deals with concrete objects	↔	Highly abstract	→
<b>Strategy</b> The extent to which the student devises (or selects) and maintains a strategy for tackling and answering the question	←	Strategy is given No need to monitor strategy  No selection of information required No organisation required	↔	Student needs to devise their own strategy Student must monitor the application of their strategy  Must select content from a large, complex pool of information Must organise how to communicate response	→

## PHASE 2 - THE INTEGRATION OF SUBJECT SPECIALISTS IDEAS

The CRAS scale was applied by experienced examiners to exam questions in History, Geography and Chemistry. It was used to gauge the demands that questions were making on students. The examiners annotated the scales for their subjects, these annotations can be used as data describing the demands residing in exam questions.

The following tables show the scale of demands as adapted by the Geography, History and Chemistry examiners.

**Figure 3 Scale of demands in Geography**

Type of Demand	1	2	3	4	5
<p><b>Complexity</b> The complexity of each component operations/ideas and the links between them.</p>	←	<p>Types of questions showing lower demands</p> <p>Description Small scale - local/personal Local/familiar Short sentences Human/ everyday</p>	↔	<p>Types of questions showing higher demands</p> <p>Analysis and synthesis Large scale - beyond personal experience Europe → Less Economically Developed Country Remote/ unfamiliar Long sentences Physical past and future</p>	→
<p><b>Resources</b> The student's use of the data/information given.</p>	←	<p>Told where to look Simple resource Designed for question Topographic/isomorphic</p>	↔	<p>Not told where to look Complex resource Not designed for question Map transformation</p>	→
<p><b>Level of Abstraction</b> The extent to which the student is required to deal with ideas rather than concrete objects or phenomena.</p>	←	<p>Everyday language Sentences Concrete - today Deals with today Idiography Locations</p>	↔	<p>Technical language Mathematical formula Imaginary - past/future Projects forward and back Pattern/distribution</p>	→
<p><b>Strategy</b> The extent to which the student is required to devise and monitor a strategy for tackling the question and organise the information to be communicated in the answer</p>	←	<p>Series of steps Deals with resources one at a time Identifies simple links between 2 ideas Makes decision with some evidence Selects case study with help/specification</p>	↔	<p>Essay with title posed and a sequence Selects from several versions Identifies several links between 3 or more ideas Makes decision, gives evidence, discusses and qualifies. Selects case study to criteria</p>	→

**Figure 4 Scale of demands in History**

	<b>1</b> ←	<b>Lower demands.</b> <b>2</b> Requires only elementary links and /or evaluation. No requirement to put two sides of an argument	<b>3</b> ↔	<b>Higher demands.</b> <b>4</b> Requires discussion of links. Focuses on evaluation. Requires balanced arguments.	<b>5</b> →
<b>Complexity</b> The complexity of each component operation or idea and the links between them.	←	Requires only elementary links and /or evaluation. No requirement to put two sides of an argument	↔	Requires discussion of links. Focuses on evaluation. Requires balanced arguments.	→
<b>Resources</b> The use of data and information.	←	Knowledge required primarily 'factual' - events, dates etc. Some inaccuracy and omissions permissible. Selection (relevance) may be obvious from question.	↔	Requires knowledge of concepts as well as 'facts'. Knowledge accurate and, where necessary, detailed. Requires developed sense of relevance.	→
<b>Abstractness</b> The extent to which the student deals with ideas rather than concrete objects or phenomena.	←	Requires narrative or description and focus on events rather than ideas. Avoids need for technical terms.	↔	Requires analysis and explanation rather than narrative or description. Requires grasp of technical terms.	→
<b>Strategy</b> The extent to which the student is required to devise (or select) and maintain a strategy for tackling and answering the question	←	Provides clear and detailed framework for answer.	↔	Requires candidate to organise own argument.	→

**Figure 5 Scale of Demands in Chemistry**

	1	2	3	4	5
<b>Complexity</b> The complexity of each component operation or idea and the links between them.	←	Simple operations (i.e. ideas/steps) No comprehension, except that required for natural language No links between operations	↔	Synthesis or evaluation of operations Requires technical comprehension Makes links between operations	→
<b>Resources</b> The use of data and information.	←	All and only the data/information needed is given	↔	Student must generate the necessary data/information.	→
<b>Abstractness</b> The extent to which the student deals with ideas rather than concrete objects or phenomena.	←	Deals with concrete objects	↔	Highly abstract	→
<b>Strategy</b> The extent to which the student is required to devise (or select) and maintain a strategy for tackling and answering the question	←	Strategy is given No need to monitor strategy No selection of information required No organisation required	↔	Student needs to devise their own strategy Student must monitor the application of their strategy Must select content from a large, complex pool of information Must organise how to communicate response	→

### **PHASE 3 - INTERVIEWS WITH EXAMINERS USING KELLY'S REPERTORY GRID TECHNIQUE.**

The Repertory Grid interview is a tool to help interviewees communicate their views and ideas using their own, meaningful language. It was developed by Kelly (e.g. 1955) within the framework of his *personal construct theory*, a theory of personality development that is widely used in research and therapy today. The primary focus of personal construct psychology was upon the way individuals perceive their environment and the way they interpret what they experience. The person is described as a 'scientist' constantly forming hypotheses in an attempt to understand the world. In the best traditions of the philosophy of science the person employs as few and as simple hypotheses as possible. Thus we can expect a person in any given context to 'make sense' of what they see by sorting and classifying all of the phenomena experienced according to a few simple criteria. These are the *constructs* of the theory, since they do not have any independent existence in the world but are built up by the person from their experience. While we may be predisposed by evolution to build certain kinds of construct rather than others Kelly emphasises the *personal* nature of the experience that goes into building them, and it is therefore clear that each person's repertoire of constructs will be different. The Repertory Grid technique is a method for making those constructs explicit. By controlled presentation of stimuli the person is provoked into revealing the criteria that are most salient for them at that moment, and from this the researcher (or therapist) can infer how the person perceives their experiences, even though the person may not be able to make the constructs explicit.

The method has been applied to education in several ways. Kremer-Hayson (1991) used it to elicit from teacher-managers their perceptions of good professional practice. Fisher et al. (1991) applied the technique to course evaluation. Parsons et al. (1983) showed how the grid technique could be used to make explicit the implicit models that teachers have of how children learn. The power of the technique is that it can make the implicit explicit, and previous research has shown that examiner's conceptions of demands tend to be implicit (SRAC 1990). Repertory grid interviews were used to enable examiners to share this tacit knowledge.

The dimensions of demand in the CRAS scale were 'imposed' on the examiners to a certain extent, although they were allowed to modify them to suit their subject. The Repertory Grid interviews allowed examiners to describe, using their own language, the constructs considered to constitute 'demands'. The purpose of the Repertory Grid analysis was to try to get at the examiners' own way of construing 'demands' in their subject - in Repertory Grid language, the 'constructs' forming their concept of demand.

#### **Procedure**

Six examiners were interviewed. Two from each subject, one GCSE specialist and one A Level specialist.

In each subject, examiners were repeatedly shown 'triads' of three questions, and asked to describe how two were similar to each other and different from the third. This process produced a set of 'constructs'. These constructs showed us what the examiners considered the demands of the questions to be. The examiners then rated each question on a scale of 1-5 for each of the constructs which they had elicited. These ratings

were factor analysed, showing us which constructs were seen as similar. This allowed us to compare the examiners' ideas of the demands in the questions with the CRAS scale and thus further adapt the scale of cognitive demands.

### Outcomes of Repertory Grid Interviews

The constructs detailed below were produced from examiners as they compared the demands in exam questions. They are presented using the examiners' language.

#### Geography GCSE

The geography GCSE examiner produced 10 constructs describing the demands in the GCSE geography questions. Some of these constructs are very similar to the demands detailed in the Geography GCSE version of the CRAS scale. Three factors emerged, suggesting that the constructs were based on three different types of demand.

Factor 1 consisted of the constructs listed below. It accounted for 53.7% of the variance in the examiner's ratings. This factor relates to the provision and use of resources and information, which is akin to the dimension 'resources' in the CRAS scale. However, factor 1 is actually broader than this; Resources in CRAS only pertained to the provision of information (constructs 4 and 6) whereas the examiner, in using the Repertory Grid technique produced other constructs which showed the importance of the use of that information (constructs 3, 5 and 8), as well as other constructs relating to what students are expected to do with the information.

#### Factor 1

3	Describe	Explain situation
4	Resource familiar Interpretation and analysis techniques familiar	Resource new or no resource provided Interpretation and analysis techniques must be developed
5	Processing simple information	Processing more complex issues
6	No recall of information necessary	Recall of information necessary
8	Provide evidence only	Provide and analyse evidence

Constructs 1 and 7 constituted the second factor (accounting for 19.9% of the variance). This factor relates to the links between different aspects of the questions. This relates very closely to the 'strategy' dimension, as well as containing the linking aspects of 'complexity'.

#### Factor 2

1	Structure provided	Structure not given
7	Links provided	Links less obvious

The third factor accounted for 13.8% of the variance. This relates to the core difficulty of the content of the geography. More demanding questions were described as involving abstract ideas, and were about remote or unfamiliar places, and dealing with issues on a large scale. More demanding questions also required that candidates make predictions. This factor is very closely related to the CRAS scale's 'abstraction'.

### Factor 3

2	More concrete	More abstract
9	Local / small scale and more familiar subject	Remote / larger scale and less familiar subject
10	Current issues Present	Prediction required

### Geography A Level

From the factor analysis of the examiner's ratings of the questions on the constructs four factors emerged, suggesting that the constructs were based on four different types of demand.

### Factor 1

4	Focus on single message from data	Focus - several different relationships to be identified
5	Indication of single unambiguous answer	Trigger word hints at uncertainty
10	Small scale, matching experience	Large scale, global, remote from experience

Factor 1 accounted for 35.4% of the variance in the examiner's ratings. In terms of the CRAS scale this factor includes a combination of complexity and abstraction.

### Factor 2

1	Straight into the question - no diagram or familiar diagram. Set steps taught	Preparation - have to understand complex diagram. Can't be pre-learned
3	Describe without explanation for a simple resource	Resource - complex/unfamiliar. Describe and explain a relationship
6	No specific variables, just description terms	Variable range - measures on different scales

Factor 2 included constructs relating to resources (accounting for 17% of the variance).

### Factor 3

7	Concrete visible processes	Abstract conceptual processes, mathematical formula
9	Contemporary processes which can be seen in the field. Everyday common experience	Involving past and future, processes no longer active

The third factor consisted of constructs such as abstraction and accounted for 13.7% of the variance in the examiner's ratings.

### Factor 4

8	Divided into series of steps	Broad statement to be explored
2	Everyday language, no technical terms	Vocabulary complex
11	Describe, define	Command words - evaluate, explain

The fourth factor included the following constructs and it accounted for 10.3% of the variance in the examiner's ratings. This factor is related to the language of the question.

### Discussion of Geography Demands

The Repertory Grid interviews and the examiners' elaborations of the CRAS scale were integrated and the new scales are shown in the next table. A number of changes were made:

We see from the fourth factor that the examiners considered language to be an influence on demands in questions (although it was not explicit in the CRAS scale). The constructs elicited from the examiners suggest that language should feature in the scale of demands for geography.

Abstraction was an important factor in the demands of the questions as the examiners saw them, and merits a dimension of its own. Much study in geography involves working in the abstract mode, for example using maps and symbols. However, it seems that the geographical content of a question can vary in abstraction in a number of ways and these ways are incorporated in the 'abstract' dimension of the new scale.

The CRAS scale described resources in very general terms, however, geography involves many different uses of different types of resources, so there is a need to be more specific about the meaning of that dimension for geography. The Scale of Demands in geography shown below incorporates the examiners' descriptions of different levels of demand in resources, and includes more about the complexity of the resources and how candidates are expected to use them. One important feature of a resource is whether or not the candidate is familiar with that type of resource, and therefore with the strategy for interpreting it. Examiners made it clear that the application of a pre-learned strategy to interpret a resource is less demanding than being presented with an unfamiliar resource which requires candidates to devise a strategy for dealing with it.

The final scale of demands for geography is shown in figure 6. It has been adapted from the CRAS scale and incorporates the examiners' perceptions of demand as generated through the repertory grid interviews.

**Figure 6 Scale of Demands in Geography Exam Questions**

	1	2 Lower demands	3	4 Higher demands	5
<p><b>Linking</b> The extent to which the question supports the candidate in answering.</p>		<ul style="list-style-type: none"> <li>• There are a number of simple steps in the question</li> <li>• Specific question focusing on one issue</li> <li>• Links between question parts are given</li> </ul>		<ul style="list-style-type: none"> <li>• The question is not broken down</li> <li>• Broad question involving several relationships</li> <li>• Candidate has to identify links between question parts</li> </ul>	
<p><b>Resources</b> The provision and use of given and internal information.</p>		<ul style="list-style-type: none"> <li>• Directed to relevant material</li> <li>• Information is readily useable</li> <li>• Resource type familiar</li> <li>• Material statistical</li> <li>• Candidate must recall/provide information</li> </ul>		<ul style="list-style-type: none"> <li>• Not directed to relevant material</li> <li>• Information needs transformation</li> <li>• Resource type unfamiliar</li> <li>• Material is linguistic</li> <li>• Information is given</li> </ul>	
<p><b>Abstraction</b> The extent to which the content of the question and its required answer are within the candidates experience.</p>		<ul style="list-style-type: none"> <li>• Time - set in present</li> <li>• Places familiar</li> <li>• Small scale</li> <li>• Individual locations/issues</li> </ul>		<ul style="list-style-type: none"> <li>• Set in past or future</li> <li>• Places unfamiliar or remote</li> <li>• Large scale</li> <li>• Patterns across locations/issues</li> </ul>	
<p><b>Processes</b> The nature of the tasks required of the candidate.</p>		<ul style="list-style-type: none"> <li>• Description/definition</li> <li>• Everyday language used and required</li> </ul>		<ul style="list-style-type: none"> <li>• Explanation/evaluation</li> <li>• Technical language used and required</li> </ul>	

**Chemistry GCSE**

The examiner found that the CRAS scale was applicable to Chemistry GCSE without further amendments, so he chose to use the scale rather than describe the demands in his own language. The ratings on the four dimensions of the scale were produced one factor, showing that the examiner had a global understanding of demand in GCSE chemistry.

### Chemistry A Level

The process of comparing triads of questions in A Level chemistry elicited a number of constructs which were factored into three dimensions

#### Factor 1

B	Familiar context (i.e. bookwork)	↔	Unfamiliar context
C	Application to unfamiliar situation	↔	Application to familiar situation

Factor 1 accounted for 41% of the variance and pertained to familiarity. In geography this would be considered to be part of abstraction (less familiar is more abstract and more demanding). However in Chemistry it was also related to how practised candidates were at particular types of questions.

#### Factor 2

A	Qualitative	↔	Quantitative
D	Information supplied	↔	Recall / selection of information

The second factor accounted for 27.8% of the variance in examiner's ratings. The Quantitative -Qualitative construct referred to the type of answer that the candidate was required to give and this affected the difficulty of questions in a similar way to construct D which referred to whether information was given or had to be recalled.

#### Factor 3

F	Organisation required	↔	No organisation required
A	Quantitative	↔	Qualitative

The third factor was labelled 'strategy' and consisted of organisation - construct F, and quantitative - construct which accounted for 20% of the variance. Construct A therefore appears as part of two different factors, so it seems that in chemistry, where the questions requires a qualitative or quantitative response affects the difficulty in terms of the strategy applies as well as the sort of knowledge brought to the question.

It is likely that the features identified as 'specificity' and 'language' have contributed to the experts' ratings of organisation too. A more general question requires candidates to decide which knowledge is appropriate and to organise that knowledge into a sentence or paragraph.

### Discussion of Chemistry Demands

Edwards' scale was originally developed for Chemistry, so it understandable that it was not amended greatly by the chemistry examiners, although it did seem better suited to GCSE than to A Level chemistry.

The final Chemistry scale of demands is shown below:

**Figure 7 A Scale of Demands in Chemistry Exam Questions**

	1	2 Lower demands	3	4 Higher demands	5
<b>Complexity</b> The complexity of each component operation or idea and the links between them.		<ul style="list-style-type: none"> <li>Simple operations (i.e. ideas/steps)</li> <li>No comprehension, expect that required for natural language</li> <li>No links between operations</li> </ul>	↔	<ul style="list-style-type: none"> <li>Synthesis or evaluation of operations</li> <li>Requires technical comprehension</li> <li>Makes links between operations</li> </ul>	
<b>Resources</b> The use of data and information.		<ul style="list-style-type: none"> <li>All data/information needed is given</li> <li>Customised information supplied</li> <li>Information in resources is simply organised/presented</li> </ul>	↔	<ul style="list-style-type: none"> <li>Student must generate the necessary data/information.</li> <li>Student must select relevant information</li> <li>Resource contains a complex pool of information.</li> </ul>	
<b>Abstractness</b> Familiarity of the material and the context in which it is set.		<ul style="list-style-type: none"> <li>Deals with concrete objects</li> <li>Familiar context (i.e. bookwork)</li> <li>Application to familiar situation</li> </ul>	↔	<ul style="list-style-type: none"> <li>Highly abstract</li> <li>Unfamiliar context</li> <li>Application to unfamiliar situation</li> </ul>	
<b>Strategy</b> The extent to which the student is required to devise (or select) and maintain a strategy for tackling and answering the question		<ul style="list-style-type: none"> <li>Strategy is given</li> <li>No need to monitor strategy</li> <li>Response strategy given</li> <li>Quantitative response required</li> </ul>	↔	<ul style="list-style-type: none"> <li>Student needs to devise their own strategy</li> <li>Student must monitor the application of their strategy</li> <li>Student must devise and organise response strategy</li> <li>Qualitative response required</li> </ul>	

### History GCSE

The examiner described a number of constructs of demand, and his ratings showed that there were two factors of demand:

#### Factor 1

	Less demanding	More demanding
1	Candidates are led through the question.	The candidate gets straight in.
3	Not necessarily asking for specific knowledge.	Asked for specific knowledge.
4	Helpers are a trigger to recall and select relevant knowledge.	Having to select relevant information for themselves.
7	Not given that opportunity.	Given the opportunity to look at both sides of the argument.
11	Everything is implicit. They have to make their own decisions about their moves and strategy.	The strategy to use is made explicit and candidate is given small chunks to work on.
12	They must manage everything.	Some issues have been raised in a previous question, so the previous question informs the final evaluation.
13	Differences and similarities are pointed out.	The candidate has to cross reference the sources for themselves.
14	They have written about every source before evaluating, so the resources have received a more detailed interrogation before evaluation.	Not interrogated all sources individually before the final evaluation. The focus of the question encouraged surface use of sources.
15	MS can't expect all moves of the structured question. Mark scheme is not an accumulation of points.	Mark scheme is explicit.

Factor 1 related to the extent to which the candidate is led through the question, being given clues about (i) how to organise and present their answer and (ii) what would be relevant material to include in their answer. This factor accounted for 72.2% of the variance in the examiner's ratings of questions and included the following constructs:

#### Factor 2

	Less demanding	More demanding
2	Never quite asked to pull the whole thing together.	Reach overall judgement.
5	Asked to do concrete/surface things. They can accumulate marks for lower level skills.	Abstract
6	Mark scheme is broader, looking for quality of answer.	Candidates have to perform pedestrian/ mundane things throughout the question. They need perseverance.
14	They have written about every source before evaluating, so the resources have received a more detailed interrogation before evaluation.	Not interrogated all sources individually before the final evaluation. The focus on sources encouraged by the question surface use of sources.

Factor 2 accounted for 19.7% of the variance in the examiners ratings and pertained to the extent to which higher order skills (e.g. evaluating, concluding and defining their own criteria for evaluation) were required. (Constructs 2, 5, 6 and negative of 14.)

GCSE History produced two quite general factors, showing that the examiner did not identify specific aspects of demand, but saw demand as one concept. This fits with the type of quality judgements that were likely to be made in history involving an overall judgement of quality.

#### History A Level

Factor analysis of the constructs showed that three factors emerged from the thirteen original constructs. These three factors accounted for 89% of the variation of the examiner's ratings.

#### Factor 1

4	The question provides structure for argument. It breaks down both sides of the argument.	The candidate has to decide for themselves how to order/provide an argument for both sides.
7	They are led through the stages of the argument.	Candidate is left on their own to identify and sort out relevant material
8	Candidates make judgements throughout and reach their own conclusion. This is a more familiar way for them to work.	Candidate needs to be trained to conclude as a separate exercise (for a few marks). Candidates are unfamiliar with this questions type.
10	They are given the freedom to structure the question as they see appropriate.	They are restricted in developing a different line of argument by the structure of the question.

Factor 1 accounted for 39.9 % of the variation in the ratings. It pertains the degree to which the students organise their answer and present their argument. The examiner clearly saw this as a very important factor contributing to the demands of these questions.

#### Factor 2

Accounting for 28.6% of the variance in ratings, factor 2 clearly relates to the identification of relevant historical material that is brought to the answer.

1	Question indicated which area to flesh-out. Ideas were	No hints to what contributions/resources
---	--	--

	suggested to them.	
2	It is easier to introduce descriptive material (e.g. events and policies).	Candidate must relate material to abstract concepts/ideas.
5	They are given the opportunity to fit material into the time available.	Candidate is asked to do too many things in 45 minutes

### Factor 3

3	The question does not overtly require evaluation. The candidate can get a fair mark from description.	Candidate must evaluate material/ideas.
11	They are given the freedom to decide how to allocate time to different arguments.	A brief answer is required to a question which could be lengthier.
12	Individual judgements are kept separate.	Required to balance complex judgements

The third factor accounted for 20.5% of variance in the examiner's ratings. This factor is about the need to link sub-questions and the difficulties candidates (who are used to responding to essay questions) have when having to respond to short answer questions.

### Discussion of History Demands

The final Scale of Demands in History is different from the CRAS scale in that it is less focused on content and more on the development of strategies for arguing and presenting answers.

**Figure 8 A Scale of Demands in History Exam Questions**

	1	2 Lower demands.	3	4 Higher demands.	5
Skills		<ul style="list-style-type: none"> <li>Separate decisions on each issues</li> <li>Thinking in abstract needed</li> </ul>		<ul style="list-style-type: none"> <li>Making an overall judgement on all material</li> <li>Surface thinking/descriptive terms</li> <li>Requires evaluation</li> </ul>	
Sources of Information		Requires description			
		<ul style="list-style-type: none"> <li>Hints to relevant material are given</li> <li>Material has been addressed in preceding questions</li> <li>Candidate must cross-reference sources</li> </ul>		<ul style="list-style-type: none"> <li>Candidate has to identify relevant material</li> <li>Material has not been introduced gradually</li> <li>Candidate deals with one source at a time</li> </ul>	
Strategy and Links		<ul style="list-style-type: none"> <li>Question is about specifics*</li> <li>Candidate has to deal with one side of argument only</li> <li>Structure for answering is provided.</li> </ul>		<ul style="list-style-type: none"> <li>Question is broad*</li> <li>Candidate has to deal with two or more sides of argument.</li> <li>Candidate needs to devise structure for answer</li> </ul>	
External influences		<ul style="list-style-type: none"> <li>Freedom⊕</li> <li>Mark scheme requirements are explicit</li> <li>Familiar question type</li> </ul>		<ul style="list-style-type: none"> <li>Restriction⊕</li> <li>Mark scheme requirements are implicit</li> <li>Unfamiliar question type</li> </ul>	

\* Which of the two ends of these demands is the more difficult is mediated by the expectation and ability of the candidate, so whether a specific or broad question is more or less demanding depends on the students expectations.

⊕ The ability of students was found to influence whether they were helped or hindered by freedom or restriction. See Pollitt et al. (1998) for a full discussion of the effects of expectation and ability on demands.

### DISCUSSION

The three scales of demands, for History, Chemistry and Geography have been developed for application of assessment tasks in these three subjects at both GCSE and A Level. A number of points for discussion have arisen:

- The scales need validation. Edwards and Dall'Alba extensively trialled the scale with school science teachers in order to validate the dimensions, as well as to test out the practicability of the scale. The subject scales, as they stand require validation. Another technical issue which needs addressing is one of inter-rater reliability the scales; this is yet to be measured.
- The three subjects studied were chosen because they incorporate many questions types (structured and essay questions) and the content of the three subjects spans most of the disciplines (mathematical, literary, and physical and social scientific). As such these three subjects are likely to provide scales which can be easily generalised to a more general scale. For researchers such a scale could provide a tool for comparing the demands made across subjects. Having said that, the scale may be generalisable to other subjects, there is a criticism of the scale which may render it less useful for some subjects, for example music or PE. It contains no affective or psychomotor demands.

### **Implications**

For the sake of validity, we should ensure that students are properly prepared for the kinds of demands they are going to meet in exam questions. The students' expectations are created by the syllabus, previous exam questions, mark schemes and INSET provided by the board. The scales can be used to see if the demands of the (i) text books and teaching materials, (ii) national curriculum (iii) lesson content (iv) assessment tasks (v) marking criteria for the assessment tasks are matched. Gagne and Briggs (1973) talk about learning objectives, learning tasks and evaluation techniques and state that 'it is necessary to achieve internal consistency amount three important components of instruction'.

The prompt for this paper was the finding that examiners find it difficult to articulate their understanding of demands in the questions they are formulating (SRAC 1990). The scales provide a language for examiners to articulate and share discussion, thus building an awareness of the demands being made on candidates as well as providing guidance on how to vary those demands.

### **REFERENCES**

- Ausubel, D. P., Novak, J. D. and Hanesien, H. (1978) *Educational Psychology: A cognitive view 2nd edition*, Hold, Reinhart and Winston USA
- Bloom, B. S. (Ed) 1956 *Taxonomy of Educational Objectives Book 1: Cognitive Domain*, McKay New York.
- de Bono, E. (1976) *Teaching Thinking*. Maurice Temple Smith, Great Britain.
- Bruner, J., Olver, R.R., Greenfield, P.M., Rigney, Hornsby J., Kenny, H.J., Maccoby, M., Modiano, N., Mosher, F.A., Olson, D.R., Potter, M.C., Reich, L.C., Mackinnon and Sonstroem, A. (1966) *Studies in Cognitive Growth*, John Wiley USA.
- 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, Melbourne, Victoria, Australia.
- Edwards, J. & Dall'Alba, G. (1981) Development of a scale of cognitive demand for analysis of printed secondary science materials. *Research in Science Education*, 11,158-170.
- Fisher, B., Russell, T. and McSweeny, P. (1991) Using personal constructs for course evaluation. *Journal of Further and Higher Education*, 15 (1) 44-57.

- Kelly, G. A. (1955) *The Psychology of Personal Constructs*. New York W.W, Norton and Company Inc.
- Kremer-Hayson, L. (1991) Personal constructs of elementary school principals in relation to teachers. *Research in Education*, 43 15-21.
- Novak, J. D. (1977) *A Theory of Education*. Cornell University Press, London.
- Parsons, J. M., Graham, N. and Honess, T. (1983) A teachers's implicit model of how children learn. *British Educational Research Journal* 9(1) 91-101.
- Pollitt, A., Hughes, S., Ahmed, A., Fisher-Hoch, H. and Bramley, T. (1998) *The effects of structuring exam questions on the demands made on students in GCSE and A Level questions*. Report to QCA/ACCAC 1998.
- Rigney, J. W. (1978) Learning Strategies: A Theoretical Persepctive. In O'Neill, H.F. Jr. (Ed) *Learning Stretegies*. New York: Academic Press.
- SRAC (1990) *A Study of the Demands Made by the Two Apporaches to 'Double Mathematics'*. An investigation conducted by the University of Cambridge Local Examinations Syndicate on behalf of the Standing Research Advisory Committee of the GCE Examining Boards.
- Taba, H. (1962) *Curriiculum Development: Theory and Practice*, Harcourt Brace & World USA
- Taba, H. (1967) *Teacher's handbook for elementary social studies* Addison-Wesley USA