Routledge
Taylor & Francis Group

Educational Research

ISSN: 0013-1881 (Print) 1469-5847 (Online) Journal homepage: www.tandfonline.com/journals/rere20

Can a picture ruin a thousand words? The effects of visual resources in exam questions

Victoria Crisp & Ezekiel Sweiry

To cite this article: Victoria Crisp & Ezekiel Sweiry (2006) Can a picture ruin a thousand words? The effects of visual resources in exam questions, Educational Research, 48:2, 139-154, DOI: 10.1080/00131880600732249

To link to this article: https://doi.org/10.1080/00131880600732249

Educational Research, Vol. 48, No. 2, June 2006, pp. 139–154

Routledge
Taylor & Francis Group

Can a picture ruin a thousand words? The effects of visual resources in exam questions

Victoria Crisp* and Ezekiel Sweiry
Cambridge Assessment, Cambridge, UK

Background
When an exam question is read, a mental representation of the task is formed in each student's mind. This processing can be affected by features such as visual resources (e.g. pictures, diagrams, photographs, tables), which can come to dominate the mental representation due to their salience.
Purpose
The aim of this research was to investigate the effects of visual resources in exam questions and, in particular, to investigate how and when students use images and whether subtle changes to these salient physical features can affect whether a question is understood and answered in the way intended by the question-setters.
Sample
The participants were 525 16-year-old students, with a range of ability, in four secondary schools.
Design and methods
Experimental test papers were constructed including six questions based on past examination questions and involving graphical elements. For five of the six questions, two versions were designed in order to investigate the effects of changes to visual resources on processing and responses. A sample of the students were interviewed afterwards.
Results
Where two versions of a question were trialled in parallel, the differences in the visual resources significantly affected marks for one question and had smaller effects on marks and the nature of answers with some of the others. There were mixed views from students over whether a visual resource that is not strictly necessary should be used. Some considered it reassuring, whilst others deemed it unnecessary. Evidence in the literature suggests that caution may be needed since there is a risk that some students may pay too much attention to the image. Findings from one question (question 6) indicated that visuals can increase the likelihood of students making unhelpful interpretations of a question. Students were seen to have sensible expectations regarding when to use information from a visual resource and what is important in an illustration. In addition, more use tended to be made of a technical diagram (in question 12) in comparison to pictures or sketches, and it was found that if an image provides a clue to an answer, this may be used in preference to information in the text.

Evidence regarding the use that students made of a table (question 1) indicated that the data in the table cells were given more attention than some of the preceding text and text in a header. This might apply similarly to other resources like graphs and charts.

*Corresponding author. Cambridge Assessment, 1 Hills Road, Cambridge CB1 2EU, UK.
Email: crisp.v@cambridgeassessment.org.uk

ISSN 0013-1881 (print)/ISSN 1469-5847 (online)/06/020139-16

DOI: 10.1080/00131880600732249

Conclusions
It is important to ensure that the inclusion of a visual resource is carefully considered and appropriately designed. If a visual resource is not strictly needed in a question, the writer will need to balance the advantages and disadvantages. Authors should also consider whether and how students are likely to use or be affected by the particular visual resource chosen. The findings and suggested implications of this study are most applicable to high-stakes testing but may also be useful to those preparing school textbooks and to teachers in their preparation of classroom materials.

Keywords: *Exam questions; Secondary school assessment; Visual resources; Diagrams; Illustrations; Pictures*

Introduction

Physical aspects of the way an exam question is presented can influence students' understanding of the task. For example, the location of question elements on a page can affect whether they are processed thoroughly and how important a particular piece of information is perceived to be. The visual resources contained in exam questions, such as pictures, tables, diagrams and photographs, can sometimes influence students' understanding and responses (Fisher-Hoch *et al.*, 1997).

The visual resources used in teaching resources (e.g. printed textbooks and teacher-prepared handouts, worksheets, tests and homework exercises) will affect students' comprehension of these materials. Hence this investigation, although set in the context of external examining, also has implications for the preparation of teaching materials.

Visual resources are sometimes included in exam questions to test students' abilities to interpret them, but they are more commonplace than this alone would warrant. There seems to have been little research into the effects of including visual resources in examination questions. However, research on the influences of illustrations in instructional texts provides some relevant insights. Such research has often found pictures to have a positive influence on learning and retention, with text being remembered better when it is illustrated (Ollerenshaw *et al.*, 1997; Weidenmann, 1989; Schnotz, 2002). However, the main purpose of exam questions is to assess learning rather than teach and hence this alone does not justify their use in exams.

Various other positive benefits reported by research on instructional texts may explain the use of visual resources in examination questions. Graphics are thought to 'simplify the complex' and 'make the abstract more concrete' (Winn, 1989, p. 127). Peeck (1993) makes a similar point when she writes that images 'might help to clarify and interpret text content that is hard to comprehend' (p. 228). It is also argued that graphics can provide more information than can be explained in words (e.g. Stewart *et al.*, 1979). These are justifiable reasons for including visual resources in exams, as including a clear illustration rather than a textual description could reduce the length of questions and help students to access questions involving abstract concepts.

In addition, images are generally believed to have a motivational role in the context of instructional texts (Peeck, 1993), which could equally apply to exam questions.

Since examinations are stressful situations for most students, elements that trigger their interest or make a question look less daunting could play an advantageous role, particularly in papers designed for less able candidates.

Not all research has found images to be beneficial. In a review of studies on instructional texts, Levie and Lentz (1982) found that in about 15% no significant effects of including images were observed. One possible explanation is that the choice of image is important. Peeck (1987), for example, found that participants who read a text without a diagram were actually more motivated and more interested to continue reading than those who read the same text accompanied by a poor diagram. This suggests that visual resources are not always beneficial and that the quality and appropriateness of a visual resource are likely to be important.

The failure of visual resources to aid instruction in some studies has often been explained as either a result of students' learning styles, as Ollerenshaw *et al.* (1997) report, or due to students not processing illustrations adequately (Weidenmann, 1989). The latter is thought to be a result of the apparent ease of processing an image giving students the false impression that they have fully understood an image when they have not (Weidenmann, 1989). In addition, Winn (1989) warns text designers of making assumptions that all students will process a particular image in a particular way. This idiosyncrasy of interpretation is also implied by Elkins (1998), an art historian, who asserts that visual images do not provide meaning via an orderly set of signs in the same way as a text.

The main risk of including images in the context of examining is that an image may lead to the formation of a mental representation of a question that does not match the meaning intended by the question-setters. When a student reads a question, a mental model (or mental representation) is constructed as a response to the text being processed (Pollitt & Ahmed, 1999). This representation is composed of images, concepts and emotions, and the relationships between concepts, but not of actual words. It is based on ideas that are already known to the reader (Johnson-Laird, 1981) and hence will be the reader's own personal understanding of the text. Therefore students' mental representations of the text may not all be the same, perhaps emphasizing certain aspects that seem particularly salient to them. Most of this process is unconscious and automatic.

Visual resources are likely to play a large role in the development of the student's mental model of a question and more emphasis is likely to be placed on the ideas communicated by them than the ideas conveyed by the associated text. As Peeck (1987) states, 'too much attention may be deployed to the illustrations themselves rather than to the accompanying text' (p. 118). She describes a previous study (Peeck, 1974), in which students were presented with a story that sometimes contained a mismatch of information between text and image. During questioning, the students tended to choose the responses consistent with the visual resources more frequently than the responses that would be indicated by the text, suggesting a dominating influence of the images.

There are a number of possible reasons for the apparent superiority of images over text. First, processing visual material may require less cognitive effort. According to Biedermann (1981), the general meaning of an image can usually be grasped in as

little as 300 milliseconds. This may be because the elements of a visual resource can usually be processed simultaneously, whereas a text must be processed sequentially (Winn, 1987).

Another perspective is that visual and textual materials may be processed in different cognitive systems. Paivio's (1975) theory of dual-coding explains the superiority of memory for images as a result of them being coded both as images and as their verbal labels whilst words are only encoded verbally. Thus the two representations of one item result in bias towards information gained from visual resources (Schnotz, 1993). Mayer (1989) argues further that the double coding of images facilitates the formation of a mental model since referential connections between the two representations will already have been produced. However, there has also been opposition to this view and some have even claimed that images provided with text might be harmful since attention is split between the two forms of information and the two sets of ideas that are triggered then have to be integrated (Sweller, 1990).

In general, placing information higher on a page will make it seem more important (Winn, 1987). However, there is also some evidence that visual resources are more likely to be 'read' and processed *before* accompanying text regardless of their relative positions. Kennedy (1974) explains how 'sometimes we read a label or caption before looking at the picture, but more often, probably, we notice the picture first and recognise the pictured object without any help from the accompanying words' (p. 7). It has been well documented that the first elements contained within a mental model will dominate and strongly influence subsequent elements (Gernsbacher, 1990). This is because a mental representation is started on the basis of the first element processed, and each subsequent piece of information is incorporated into the developing representation whenever possible. Hence the fact that visual resources are likely to be processed first means they will be likely to dominate the representation.

If visual resources do have a disproportionately large influence on the development of mental models, this has strong implications in examinations where students' ability to process material effectively is already compromised by test anxiety (Sarason, 1988). Students need to understand questions in the way intended in order to have a fair opportunity to display their knowledge and skills.

Method

A total of 525 students, aged 16 years, from across four secondary schools sat an experimental science test under exam conditions. The test included six questions involving graphical elements. For most of these questions, two versions were constructed in order to investigate the effects of changes to visual resources on processing and responses. The six questions, along with some additional common questions, were compiled into two versions of a test paper that were assigned to students at random. Twenty-seven pairs of students were interviewed after they had taken the test.

The predicted GCSE grades of the students ranged from A* to G and their distribution was fairly typical of the national school population. Predicted grades were converted into scores from 8 points for a grade A* to 1 point for a grade G. The mean

score for students attempting the test was 4.50 ($N = 261$, $SD = 1.369$) for version 1 of the test and 4.55 ($N = 254$, $SD = 1.353$) for version 2, suggesting that the two groups were very similar in ability. The maximum mark available on the test was 44 and the marks achieved ranged widely (from 5 to 41). The average mark achieved was 26.25. The correlation between the predicted grade scores and the test marks was $r = 0.63$. This suggests that the experimental test discriminated between pupils fairly well.

Results

Question 1: Insects

In version 1 of this question, the sentence 'All numbers are in thousands' was positioned above the table. This format is common in tables published in books and magazines. The layout of the various elements on the page is such that even if students started to read the text first, their attention might then turn to the table due to the reference to it in the first sentence of the text. As a result, the second sentence may not be read or sufficiently processed. In including the same information within the table in version 2 of the question, it was thought that it might be less readily overlooked due to its closer proximity to the numbers to which it refers.

Only answers of '170,000' and not '170' were given credit. Such a mark scheme might be considered harsh since the latter of these answers shows that a student has located the appropriate piece of information. However, such a mark scheme could be justified where the aim of the question is to interpret tabulated data accurately. In version 1, only 31.2% of students wrote the accepted answer of '170,000': 65.0% of students in this version (almost all of those who did not gain the mark) wrote '170'. There was a slight improvement with version 2 of the question where 33.6% of students gained the mark, whilst 63.7% answered '170'; however, this is a small difference that is not statistically significant. It seems that most students focused exclusively on the figures contained in the table when answering and tended not to refer back to other details provided in the text. Even placing the information about thousands within the table header had only a small impact, if any.

Version 1

1 The chart below shows the average number of insects present per 10 km² in a woodland area over a number of months.

All numbers are in thousands.

Name of Insect Species present in Woodland Area	May	June	July	August
Greenfly	110	120	170	200
Wasp	15	16	15	16
Ladybird	6	5	3	2
Dung Beetle	10	11	11	11

Use the information above to answer the following questions.

(a) What was the average number of greenfly present per 10 km² in July?

.....[1]

Version 2

1 The chart below shows the average number of insects present per 10 km² in a woodland area over a number of months.

	Number of insects, in thousands			
Name of Insect Species present in Woodland Area	May	June	July	August
Greenfly	110	120	170	200
Wasp	15	16	15	16
Ladybird	6	5	3	2
Dung Beetle	10	11	11	11

Use the information above to answer the following questions.

(a) What was the average number of greenfly present per 10 km² in July?

.....[1]

Question 1. Insects

From interviewing the students, it became apparent that most of those who answered incorrectly had not noticed the information that the numbers were in thousands. Indeed, most were startled to discover that their answer was incorrect, several accusing the question of being 'sneaky'.

Some students read only the first part of the text thoroughly and skimmed over the crucial piece of information, regardless of its location. One student said:

> I just read that bit [first sentence] and then looked at the chart and then looked at the question.

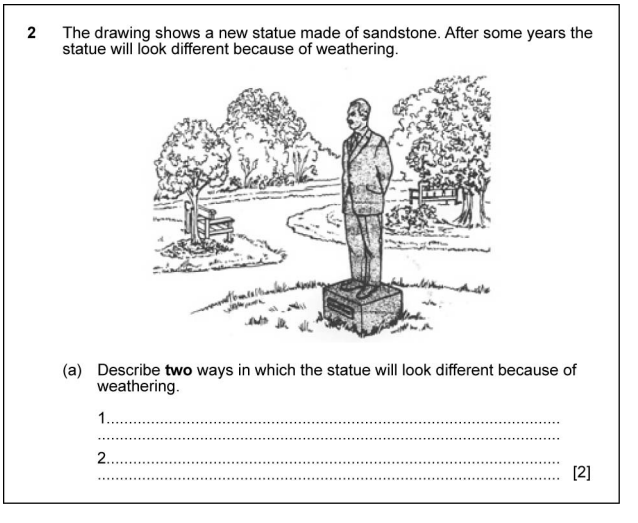
There were also some students who had processed the information, but did not write the full answer because they did not think it necessary to do so when the numbers in the table were not written out in thousands. One student said:

> I just wrote down the answer that was in the box because I didn't think it was up to me to write it down because it's already said that it was in thousands.

The text containing the vital information appears to have been low in salience. Even when students did read the information, some did not expect that this information was required when answering, given how the values in the table were presented.

Question 2: Statue

This question was common to both versions of the test and was included to investigate what use students make of a diagram that is not essential. Answers such as 'discoloured', 'material worn away' and 'bits broken off' were credited. Responses



Question 2. Statue

referring to a named feature such as nose or eyes being altered were also accepted. Hence, students did not necessarily need to use the diagram but it could help them to gain one of the marking points. The picture could also potentially fulfil a motivational role and make the context of the question more concrete. Students tended to score well on this question. Some 85.5% of students scored both marks, whilst 12.4% scored 1 mark and only 2.1% scored none.

There was a mixture of responses during interviews with regard to the picture. Some said they found it useful, with slightly fewer stating that they did not find it of use or that it was unnecessary. Positive comments included some impression of reassurance at having an illustration, and mention of how they could see that the statue had detailed features that might be lost. For example, one said:

> It gave you an idea of what to pick out . . . like you would lose a lot of detail on the person.

Some students held quite strong views that the picture was unnecessary. For example, one claimed that:

> You don't even need the picture. I mean if you say a statue it could be of anything, it works the same way.

There is also a possibility that the use of a picture may affect whether students use scientific reasoning. One student commented that:

> I think if I didn't have the diagram there, I would probably use more science instead of just saying its features won't be so defined. I would probably have said acid rain would sort of wash it away.

Another interesting comment supported the view that the inclusion of an illustration could reduce the amount of attention paid to the text:

> It might lead you to look at the picture instead of the text, so the answer might be wrong because you haven't read the text properly.

Perhaps this risk should be considered when deciding whether to include a visual resource when it is not absolutely necessary. The students' varied comments about the image in question also emphasize the point made earlier that not all observers will use an image in the same way.

Question 5: Children's meal

Version 1 of question 5 acted as a control. In the visual resource for version 2, an unrealistically large portion of chicken nuggets was illustrated in order to investigate whether the salience of this might dominate students' thinking and lead to answers about overeating.

Version 1

5 Use the information below to help you answer the following question.

Children's meal		
Chicken nuggets		Chicken, wheat flour, maize flour, hydrogenated vegetable oil, salt, modified starch. Raising agents, mono calcium phosphate, sodium bicarbonate, sodium aluminium phosphate. Starch, spices, whey powder, pepper, dextrose, vegetanbe oil. Acidity regulator, calcium lactate. Emulsifiers, phosphate salt.
French Fries		Potatoes (cooked in our own vegetable oil) dextrose, Salt
Milkshake		Milk, skimmed milk, cream, sugar, skimmed milk powder, glucose. Stabiliser, guar gum, sodium polyphosphate, carrageenan and carboxymethylcellulose. Vanilla flavour.

Give **two** reasons why it would not be advisable for a child to eat this meal every day.

1.....

.....

2.....

..... [2]

Version 2

5 Use the information below to help you answer the following question.

Children's meal		
Chicken nuggets		Chicken, wheat flour, maize flour, hydrogenated vegetable oil, salt, modified starch. Raising agents, mono calcium phosphate, sodium bicarbonate, sodium aluminium phosphate. Starch, spices, whey powder, pepper, dextrose, vegetanbe oil. Acidity regulator, calcium lactate. Emulsifiers, phosphate salt.
French Fries		Potatoes (cooked in our own vegetable oil) dextrose, Salt
Milkshake		Milk, skimmed milk, cream, sugar, skimmed milk powder, glucose. Stabiliser, guar gum, sodium polyphosphate, carrageenan and carboxymethylcellulose. Vanilla flavour.

Give **two** reasons why it would not be advisable for a child to eat this meal every day.

1.....

.....

2.....

..... [2]

Question 5. Children's meal

There were no responses about portion size or about overeating with either version of the question. The performance of students was very similar on the two versions. On version 1, 23.8% of students scored 1 mark and 72.8% scored 2 marks, whilst on version 2, 22.7% scored 1 mark and 72.3% scored 2 marks. On average, students did very slightly better on version 2 but this was not a statistically significant difference and seems unlikely to have been due to the visual resource.

The illustration was not as salient as was predicted. Most of the interviewees who had attempted version 2 had not noticed the large portion size in the diagram. In addition, most comments suggested that students were not viewing these images as an actual meal but as generic illustrations of the food types. For example, after the portion size had been mentioned by the interviewer, one student said:

> It doesn't really matter though, it just shows, illustrates what it is but it doesn't really matter how much there is.

Several students made comments along the lines of 'I didn't really look at the pictures, I went straight to the ingredients'. Additionally, the students who did notice the portion size did not use this in answering, perhaps not expecting an answer relating to quantity of food to be worthy of credit in science. One student commented that:

> You're thinking about it as in science, so it would be like the content in it not the amount.

One student made an interesting comment in favour of including images in exam questions:

> The use of pictures isn't particularly useful in trying to answer the question, but it's quite daunting on the day if all you've got is text and you've just got to read it, so maybe a picture would calm your nerves.

Students, in this question, were good at knowing that they should not place emphasis on the information in the picture. It would be interesting to know whether students would have made more use of the picture if it had been less generic in nature (e.g. if it had shown a child eating the meal).

Question 6: Products

Some words can trigger more than one meaning in students' minds. The word 'products' has a specific meaning in chemistry but a more familiar meaning for students might be that of 'household products'. Both are likely to be triggered in students' minds with question 6, although students with sufficient subject knowledge are likely to be able to suppress and ignore the irrelevant idea. The aim of including this question was to investigate whether the use of an image of some household products (in version 1) could affect the interpretation of the word 'products' by leading students' thought processes away from the chemical meaning of the word. This would, of course, not be desirable in an exam, but could occur by accident if care was not taken when choosing an image. Version 2 acted as a more neutral control question.

Of the students who answered version 1, 29.2% scored 1 mark, 16.4% scored 2 marks and 6.6% scored 3 marks. The same figures were 37.1%, 15.7% and 6.1% respectively for version 2. Some 6.8% fewer students scored at least one mark in version 1 than in version 2. These differences were not statistically significant. However, in version 1, 9% of students gave answers such as 'shampoo' and 'soap' (words that are assumed to have been activated by the word 'product' combined with the presence of the photograph), whilst only 1.5% of students gave such answers in version 2. The inclusion of the photograph showing household products in version 1

Version 1

6 After eating a meal, your mouth becomes very acidic. This acid can damage your teeth.

[Part (a) omitted]

Brushing your teeth with toothpaste will neutralise the acid. This will protect your teeth from damage.

[Part (b) omitted]

Some brands of toothpaste contain sodium carbonate.

(c) Three products are made when sodium carbonate reacts with hydrochloric acid.

What are they?

1.....

2.....

3..... [3]

Version 2

6 The paper in modern books contains slight traces of acid. The acid in the paper can make it slowly decay.

[Part (a) omitted]

One method of neutralising the acid in books is to use sodium carbonate.

[Part (b) omitted]

(c) Three products are made when sodium carbonate reacts with hydrochloric acid.

What are they?

1.....

2.....

3..... [3]

Question 6. Products

seemed to increase the likelihood of the inappropriate meaning of the word 'products' dominating students' mental models of the question.

Obviously students need to be able to cope with the chemical meaning of 'products' in science and no credit can be given to students using the everyday meaning of a word in a case like this. Whilst it is likely that students who wrote answers along the lines of 'soap' were those with less solid science knowledge and skills, it is important that the visual resources used or the context chosen are not such that they make inappropriate errors more likely.

Question 9: Components of a balanced diet

This question requires students to name the two other components of a balanced diet, other than 'protein'. Version 2 acted as a control. Version 1 used corned beef as an example and included a photograph to see whether this would make students more likely to answer with example *foods* rather than *components* (e.g. 'potatoes' rather than 'carbohydrates').

There was actually very little difference between responses in the two versions, both with respect to actual marks and the extent to which students wrote names of foods rather than components. Students attempting version 1 of the question were very slightly more likely to score 1 mark (52.9% vs 50.8%) and 2 marks (35.6% vs 33.3%) but these differences are small. The apparent slight advantage of version 1 could be a result of the example of corned beef helping to remind students of what protein is and hence helping them to recall other components, perhaps by thinking of examples first. There was no significant difference in the number of students who gave names of foods rather than components as their answers (1.5% of students in version 1 and 2.7% of students in version 2), suggesting the photograph of the corned beef had little effect in this respect. Whilst comments from the interviewees clearly showed that the photograph was attended to, students were also aware that it was irrelevant. One student remarked that:

> The picture doesn't do anything ... it's not really useful for the question.

Version 1

9 A meat canning company advertises its corned beef as high protein and easy to digest.

(a) Corned beef is rich in protein which is one of the major components of a balanced diet.

Name the other **major** components of a balanced diet.

1.....

2..... [2]

Version 2

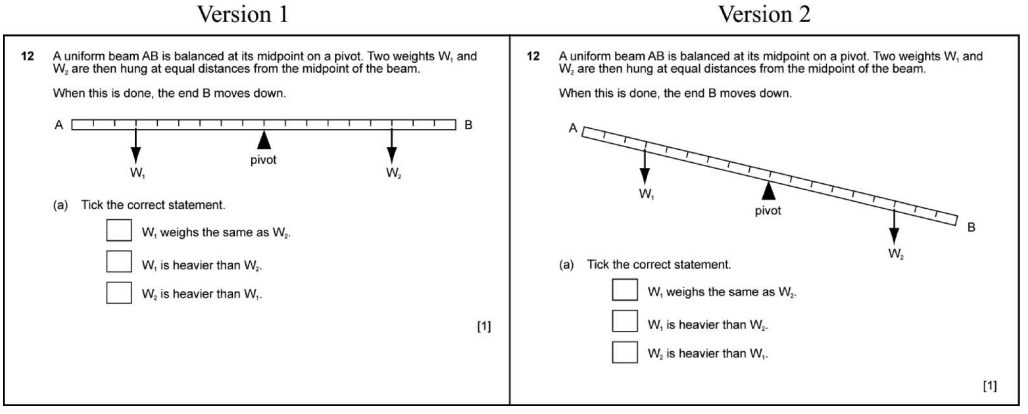
9 (a) Protein is one of the major components of a balanced diet.

Name the other **major** components of a balanced diet.

1.....

2..... [2]

Question 9. Components of a balanced diet



Question 12. Balance

Question 12: Balance

In version 1 of question 12, the text states that end B moves down when weights are attached to the beam, but the beam is horizontal in the diagram. In version 2, the beam's realistic position is illustrated. The aim was to investigate whether students attempting version 1 might overlook the textual information since the diagram appears to be able to supply the answer.

There was a statistically significant difference between marks on the two versions of the question (80.4% in version 1 and 98.8% in version 2 scored the mark, $t = 7.20$, $p < 0.001$). Some 16% of the students taking version 1 answered that the two weights were equal, suggesting that they had paid more attention to the diagram than the last sentence of the introduction.

Some students clearly expected the diagram to reflect the answer. One student said:

> Oh that actually says it moves down but it hasn't got it in the picture. That confused me because it's got the text saying one thing and the picture saying they're level.

In a similar vein, another said:

> I just didn't read the last bit or I might have read it and then ignored it and then I looked at the diagram and then went from the diagram.

Discussion

One limitation of this research is the small number of questions investigated, which has allowed investigation of only a subset of issues relating to the use of visual resources in exam questions. Unfortunately, investigating a large number of questions can be difficult as such research relies on the cooperation of school staff. School timetables and ethics place a limit on the maximum duration of a test and the good practice of including common questions within experimental test papers provides an extra

constraint. Despite the small number of questions investigated and the statistically insignificant differences for four of the five versioned questions, the research does provide insight beyond the specific examples involved. In addition, where findings from this study are in keeping with research literature from the context of instructional texts, the generalizability of findings is improved and the relevance of research on instructional texts to the context of exam questions is to some extent confirmed.

First, question 2 (statue) as well as some student comments with regard to other questions (question 5, children's meal, and question 9, components of a balanced diet) provide evidence to inform decisions over whether to include a visual resource, particularly when it is not strictly needed. Some students found having an illustration reassuring and implied that they made a question seem less daunting. Others viewed such images as unnecessary and not useful. One explanation for the diversity might be that student ability affects how students use illustrations and hence whether students view them as helpful. Those with weaker reading skills are likely to lack confidence in their ability to process text and hence generally refer to illustrations more often in search of clues (Willows *et al.*, 1981), whilst those with better reading skills are likely to refer to illustrations less and hence view those that are not strictly needed as redundant. Unfortunately, weaker readers may be exactly those who are most likely to use images inappropriately, for example, being less skilled at distinguishing between relevant and irrelevant information and at integrating information from pictures and text (Peeck, 1993).

Since in this study there were slightly more positive than negative comments and most negative comments considered visual resources irrelevant rather than detrimental, it could be argued that there is no harm in including a visual resource even when it is not strictly necessary. However, consideration must also be given to the risk that some students might pay too much attention to the image at the expense of information in the text or might not use the image appropriately. According to Peeck (1993), evidence has shown that in the context of instructional texts, 'Learners will be more likely to pay attention to pictures, and illustrations are more likely to have beneficial effects, when the text material is difficult to comprehend, and not readily understood without a picture' (p. 230), or when learners are unfamiliar with the content described by the text. Further to this, there is evidence that when visuals are attended to, the likelihood of benefit depends on the correspondence of the picture with the intended meaning of the text (Goldstein & Underwood, 1981) and that incomplete or uninformative illustrations may be detrimental to the novice learner compared to text alone (Ollerenshaw *et al.*, 1997). As the statue picture does include some useful information, perhaps its presence can be justified on that basis; however, in general it is likely that visual resources are of most use when they serve a more specific purpose such as clarifying difficult or novel content or supplying data, and that where visuals are included, their clarity, completeness and correspondence with the text are crucial. Additionally, the findings from question 6 (products) indicate that it may be beneficial to be aware that visuals can increase the likelihood of students making unhelpful interpretations of a question.

The question involving the balance (question 12) was the only one for which there was a statistically significant difference between performance on the two versions of

the question. The visual resource in this question is of a slightly different nature to those in the other questions: first, it is a more technical diagram, and secondly, the diagram in version 2 provides a strong clue to the answer. Taking the second of these points first, in comparing the two question versions it is not the case that version 2 provides more information than version 1 as both versions indicate that end B moves down. However, version 2 indicates this in the diagram as well as in the text, whilst version 1 provides inconsistent information between text and diagram. In version 1 the diagram seems to have led some students away from the correct answer due to its relative salience despite the clarity of the information in the text. The diagram in version 2 appears to have clarified the information for a significant percentage of students. This evidence is consistent with the view of Goldstein and Underwood (1981) and the findings of Peeck (1974).

A further reason for the bigger impact on students' behaviour in this question than in the others may be the nature of the visual resource. The resource in the balance question is a more technical diagram whilst those in the other questions researched in the study are less technical pictures, photographs or sketches. It seems that more technical or more diagrammatic visuals are likely to be perceived as more appropriate to use in answering than less technical images. This may be a result of the expectations that students have developed as a result of their past experiences of tests, exams papers and so on (Sweiry *et al.*, 2002). It could be argued that as well as affecting the amount that students use visuals, the nature of a visual can influence the type of reasoning that students use when attempting a question. Evans and Over (1996) distinguish between two kinds of reasoning, what can be called 'naturalistic reasoning' which is innate and is used in everyday functioning, and 'formal reasoning' which is logical and learnt. It is the latter reasoning that exams seek to assess. There is some evidence from students' comments about question 2 (statue question) which would suggest that scientific diagrams are more likely to encourage formal reasoning, and naturalistic pictures are more likely to elicit inappropriate naturalistic reasoning.

The research findings also provide evidence that students have sensible expectations regarding what is important in a visual resource. For example, the students realized that in the visual resource for the children's meal question (question 5) it was the ingredients that were important, that the images were intended to be generic rather than showing the actual meal, and that the size of portion shown was not relevant to answering. Students also realized that the image of the can of corned beef should not be used in question 6. It seems that students' expectations can prevent a potentially salient feature of an image (e.g. the large portion size of chicken nuggets) from dominating their ideas.

Question 1 (insects) gives us some insights into how tables are used by candidates in an exam. It seems that the data in the cells were the focus of students' attention. This caused a lack of attention to the text preceding the table and even moving important details into the table caption still resulted in it not being used by a significant number of students. The complication in this particular case was that even where the detail about numbers being in thousands was processed by students, it was sometimes considered unnecessary because of the way the figures were presented in the table cells. These sorts of issues may apply equally to graphs, charts, maps and

even diagrams. For example, the elements that are part of the data displayed in the resource, such as the bars or plotted points of a graph, may receive more attention in students' thoughts to the detriment of surrounding text, headings, labels and map legends. This was also seen in version 1 of the balance question where the diagram was more salient and seemed to overwrite the preceding text for some students. This relates to the finding from the context of instructional texts that information gained from a diagram will dominate in comparison to information gained from text (e.g. Schnotz, 1993).

Conclusion

The use of visual resources in examination papers can serve various positive purposes. However, the effects of images are somewhat unpredictable and hence caution is required. It is important to ensure that, when used, visual resources are accurate and complete, contain minimal irrelevance and do not cause ambiguity. If a visual resource is not strictly needed in a question, the writer will need to balance the advantage that it may make the question seem less daunting against the possible risks that parts of the text may not be read thoroughly or that a student may be led astray by an element in a visual resource that was not intended to be important.

Considering the salience and perceived relevance of visual elements can aid the prediction of their effects. The relative salience (or prominence) of visual elements can affect the extent to which they are attended to and used (e.g. data within a table may be given more attention than other text). Therefore the elements that are key to the question should be the most salient ones in the question as a whole, whether they are presented in text or by illustration. If the key information is in the text, then any visual resource should support it rather than contradict or draw attention away from it. In addition, it is worth question-setters considering students' expectations about visual resources since they can affect which elements of a question students perceive as relevant (e.g. information provided in a diagram may receive more attention than information from a sketch).

Although using visual resources appropriately is probably most crucial in high-stakes testing where there is no teacher to mediate students' understanding of a question, it is also important in teaching materials as any misunderstandings of concepts could be long-lasting. As a result, the findings and suggested implications of this study may also be useful to those preparing school textbooks and to teachers in their preparation of handouts and internally set tasks or homework exercises. Such advice for teachers may become increasingly important if teacher assessment comes to play a greater part in summative assessment (e.g. at key stage 3) as a result of the Tomlinson Report (2004). Ensuring that teacher assessment is valid and fair would become of crucial importance and assurances that what teachers think they are testing in the classroom is what is actually being tested would be required.

This study constitutes a further stage in the collection of empirical evidence on the effects of features of exam questions on difficulty and validity. The information obtained from such research is used to inform training for question writers.

Acknowledgements

The question referred to as 'Question 2: Statue', in this paper, is reproduced by permission of QCA. The question referred to as 'Question 5: Children's meal' is adapted from GCSE Design and Technology: Food Technology 1460 paper 3, 1998, and is reproduced with the kind permission of OCR. The questions referred to as 'Question 6: Products' and 'Question 9: Components of a balanced diet' are adapted from GCSE Salters Science Double Award 1774 paper 1, 2000, and are reproduced with the kind permission of OCR. The question referred to as 'Question 12: Balance' is adapted from IGCSE Physics 0625 paper 2, 1999, and is reproduced with the kind permission of CIE (University of Cambridge International Examinations).

We should also like to thank the teachers and students at the schools involved for their participation.

References

Biedermann, I. (1981) On the semantics of a glance at a scene, in: M. Kubovy & J. R. Pomerantz (Eds) *Perceptual organization* (Hillsdale, NJ, Erlbaum).

Elkins, J. (1998) *On pictures and the words that fail them* (Cambridge, Cambridge University Press).

Evans, J. St. B. T. & Over, D. E. (1996) *Rationality and reasoning* (Hove, Lawrence Erlbaum).

Fisher-Hoch, H., Hughes, S. & Bramley, T. (1997) *What makes GCSE examination questions difficult? Outcomes of manipulating difficulty of GCSE questions: a paper presented at the British Educational Research Association Annual Conference* (Cambridge, University of Cambridge Local Examinations Syndicate). Available online at: http://www.cambridgeassessment.org.uk/research/confproceedingsetc/ (accessed 14 April 2006).

Gernsbacher, M. A. (1990) *Language comprehension as structure building* (Hillsdale, NJ, Erlbaum).

Goldstein, R. & Underwood, G. (1981) The influence of pictures on the derivation of meaning from children's reading materials, *Journal of Research in Reading*, 4, 6–16.

Johnson-Laird, P. (1981) Mental models of meaning, in: A. K. Joshi, B. L. Webber & I. A. Sag (Eds) *Elements of discourse understanding* (Cambridge, Cambridge University Press).

Kennedy, J. M. (1974) *Psychology of picture perception* (San Francisco, CA, Jossey-Bass).

Levie, H. W. & Lentz, R. (1982) Effects of text illustrations: a review of research, *Educational Communication and Technology Journal*, 30, 195–232.

Mayer, R. E. (1989) Systematic thinking fostered by illustrations in scientific text, *Journal of Educational Psychology*, 81(2), 240–246.

Ollerenshaw, A., Aidman, E. & Kidd, G. (1997) Is an illustration always worth ten thousand words? Effects of prior knowledge, learning style and multimedia illustrations on text comprehension, *International Journal of Instructional Media*, 24(3), 227–238.

Paivio, A. (1975) Imagery and long-term memory, in: A. Kennedy & A. Wilkes (Eds) *Studies in long-term memory* (London, Wiley).

Peeck, J. (1974) Retention of pictorial and verbal content of a text with illustrations, *Journal of Educational Psychology*, 66, 880–888.

Peeck, J. (1987) The role of illustrations in processing and remembering illustrated text, in: D. M. Willows & H. A. Houghton (Eds) *The psychology of illustration, Vol. 1: Basic research* (New York, Springer-Verlag).

Peeck, J. (1993) Increasing picture effects in learning from illustrated text, *Learning and Instruction*, 3, 227–238.

Pollitt, A. & Ahmed, A. (1999) A new model of the question answering process. Paper presented at the *International Association for Educational Assessment Annual Conference* (Cambridge, University of Cambridge Local Examinations Syndicate). Available online at: http://www.cambridgeassessment.org.uk/research/confproceedingsetc/ (accessed 14 April 2006).

Sarason, I. G. (1988) Anxiety, self-preoccupation and attention, *Anxiety Research*, 1, 3–8.

Schnotz, W. (1993) Some remarks on the commentary on the relation of dual coding and mental models in graphic comprehension, *Learning and Instruction*, 3, 247–249.

Schnotz, W. (2002) Towards an integrated view of learning from text and visual displays, *Educational Psychology Review*, 14(1), 101–120.

Stewart, J. H., Van Kirk, J. & Rowell, R. (1979) Concept maps: a tool for use in biology teaching, *American Biology Teacher*, 41, 171–175.

Sweiry, E., Crisp, V., Ahmed, A. & Pollitt, A. (2002) *Tales of the expected: the influence of students' expectations on exam validity: a paper presented at the British Educational Research Association Annual Conference* (Cambridge, University of Cambridge Local Examinations Syndicate). Available online at: http://www.cambridgeassessment.org.uk/research/confproceedingsetc/ (accessed 14 April 2006).

Sweller, J. (1990) Cognitive processes and instructional procedures, *Australian Journal of Education*, 34(2), 125–130.

Tomlinson Report. Department for Education and Skills. Working Group on 14–19 Reform (2004) *14–19 curriculum and qualifications reform: final report of the Working Group on 14–19 Reform* (London, DfES).

Weidenmann, B. (1989) When good pictures fail: an information-processing approach to the effect of illustrations, in: H. Mandl & J. R. Levin (Eds) *Knowledge acquisition from text and pictures: Advances in Psychology 58* (Amsterdam, Elsevier).

Willows, D. M., Borwick, D. & Hayvren, M. (1981) The content of school readers, in: T. G. Waller & G. E. Mackinnon (Eds) *Reading research: advances in theory and practice, Vol. 2* (New York, Academic Press).

Winn, W. (1987) Charts, graphs, and diagrams in educational materials, in: D. M. Willows & H. A. Houghton (Eds) *The psychology of illustration, Vol. 1: Basic research* (New York, Springer-Verlag).

Winn, W. (1989) The design and use of instructional graphics, in: H. Mandl & J. R. Levin (Eds) *Knowledge acquisition from text and pictures: Advances in Psychology 58* (Amsterdam, Elsevier).