

THINKING SKILLS

Paper 9694/21
Critical Thinking

Key messages

Teachers and candidates should use old mark schemes and Principal Examiner's Reports in order to familiarise themselves with the kinds of answers expected to types of question which frequently occur. For example, candidates need to know what is meant when they are asked to identify an unstated assumption or to judge the reliability of evidence, because if they do not understand what they are being asked to do they are highly unlikely to succeed in doing it.

The number of marks allocated to a question is a good guide to the complexity and length of answer expected. Candidates often achieve only 1 mark out of 3 because they make only one point when several are available or do not develop that point in a way that would earn the additional marks.

Questions instructing candidates to 'explain the significance' do not mean the same as those asking 'how significant'. The former wording is used only when the evidence **is** significant and this type of question does not require candidates to make a judgement. Since the Examiners have committed themselves to the belief that the evidence in such cases is significant, it is not a good strategy for candidates to argue that it is not.

General comments

There was a wide range of performance on this examination. There was, as usual, a significant number of candidates who did not appear to know what they were expected to do, but at the other extreme a very small number of candidates performed exceptionally well.

Comments on specific questions

Question 1

Candidates engaged well with the sources and the story.

- (a) There were many ways of achieving the marks available for this question, and a fair proportion of candidates scored 1 out of 2, but not many were awarded 2 marks. Candidates who scored 0 tended to repeat the content of the passage rather than making a link with the suspension of JR. In **Question 1 (a) and (b)** the expression 'biased towards' was interpreted as 'biased against', which appears to be what candidates intended.
- (b) Many candidates achieved 1 mark out of 3 by pointing out that the reliability of the source was compromised by FB's bias against JR, and a fair number recognised that the headteacher had not witnessed the event himself and was relying only on FB's testimony, or that he had failed to seek corroboration of FB's account. A number of candidates claimed that only FB had a problem with JR, which – although it is true of Source D taken in isolation – overlooks the claim in Source C that FB acted 'at the request of colleagues'.
- (c) (i) The usefulness of evidence comprises reliability and significance, and so answers of either kind were acceptable. Many candidates recognised that the evidence was strengthened because AL was an eye-witness of what occurred. Several other correct answers were available, but few candidates achieved 2 marks. Those who scored 0 either summarised the content of AL's evidence instead of showing why it was useful or made evaluative points based on speculation.

- (ii) A significant minority of candidates rightly suggested that AL's embarrassment or resentment at FB's interference in her class may have led her to exaggerate the inappropriateness of his behaviour. Marks were not awarded for speculative answers, such as that AL may have had a pre-existing grudge against FB or may have been a friend of JR's family. An interpretation of the words with which she opened her memo ('I believe the incident...') as indicating uncertainty on her part that this was indeed the same incident referred to in Source D was based on a misunderstanding and not credited.
- (d) Most, but by no means all, candidates judged that JR did not deserve to be suspended. Those who thought his suspension was deserved tended to argue that swearing at a teacher is always a serious offence, even when provoked. Others, however, considered that failing to do homework, absenting oneself from detention and swearing at a teacher are relatively minor offences, and that suspension should be reserved for candidates who damage school property or bring contraband onto the premises. A few candidates were not given a mark for a conclusion because they declined to come to a judgement, gave two contradictory conclusions or made a significantly different judgement from the one stated in the question, e.g. that JR did deserve to be punished. Many candidates made good use of inferential reasoning from the sources. Some candidates considered and rejected an alternative judgement, but others misinterpreted this part of the task and suggested alternative disciplinary actions or how circumstances might have been different.

Question 2

- (a) Nearly all candidates correctly judged that the blogger had supported her claim badly, although a few made more positive judgements, such as 'somewhat supports', which were not credited. There were several ways of gaining the additional two marks and a fair proportion of candidates identified one or even two of them. Some other answers were judged to be too marginal for credit although valid as justification for the judgement mark. One of these was that with only 2 balls, there is always a 50 per cent chance of the dog bringing the right ball first time; this ignores the claim that the dog 'almost always' gets it right. The criticism that that the blogger was contradicting what 'Scientists and vets claim' was not credited as a weakness in her own evidence. A few candidates interpreted this as a question about reliability, commenting that the blogger was probably biased in favour of her own dog, but that is not what was asked.
- (b) The relatively few candidates who understood what was meant by a factor which increases or reduces the reliability of the evidence tended to observe (correctly) that reliability of the survey was increased by the owners' expertise or ability to see the behaviour of their dogs and reduced by their bias in favour of their dogs, which might lead them to exaggerate their abilities. Answers which were not credited included speculative comments about the size of the sample and the claim that the findings were reliable because they were based on research. Some candidates thought that Source C was showing how dogs' intelligence varies with age (of the dog), while others thought the percentages referred to the fraction of the intelligence of a human rather than the proportion of respondents giving the answer, e.g. 'Dogs have 38 per cent of the intelligence of 1–3 year olds.'
- (c) Although the wording of the question was intended to advise candidates that the source was significant, a significant proportion argued that it was not and were awarded 0 marks. Many candidates achieved 1 mark by suggesting that – like Hans, the horse – dogs, too, might pick up cues from their handler. Some gained a further mark by suggesting that this was a form of intelligence. There were several ways of scoring the third mark, but not many candidates did so. A significant proportion of candidates accused Hans and his owner of perpetrating trickery or a hoax, even though the source makes it clear that the misrepresentation was unintentional, while some other candidates accepted the claim that Hans was able 'to perform arithmetical calculations and to understand German', even though the source makes it clear that this was proved to be untrue.
- (d) As on previous occasions, many candidates accurately summarised the sources and drew an acceptable conclusion, thereby achieving 2 or 3 marks out of 6, while others accessed the higher marks, by making good use of evaluation and/or inferential reasoning. It was not easy to make use of Source D (although some candidates managed to do so) and many candidates misunderstood or ignored at least one other source, thereby limiting their mark for use of sources. A significant number of candidates gained a mark for personal thinking, based on their own experience as dog owners. A few candidates discussed whether dogs were more intelligent than humans, which was not the meaning of the claim.

Question 3

- (a) A fair proportion of candidates correctly identified the sentence in which the conclusion was located, but most of them included all of the sentence in their answer instead of just the latter part and thereby scored 1 mark instead of 2. The most popular wrong answers were the final sentences of the first and last paragraphs, which suggests that some candidates were answering on the basis of location rather than analysing the structure of the reasoning.
- (b) Most candidates correctly identified two or three intermediate conclusions. Many of those who recognised that the second sentence in paragraph 2 included an intermediate conclusion cited the whole sentence, including two additional elements, which was not credited. The most popular wrong answer was the first sentence of paragraph 4.
- (c) A fair number of candidates correctly identified the appeal to tradition in paragraph 2, and some of them gained 2 marks by explaining why it was invalid. Some candidates correctly identified the assumption that expensive lessons are required to pass a driving test, while two of the most common answers for a single mark were challenges to this assumption, namely that driving lessons can be obtained free of charge at school or *via* friends and family or that it is possible (in some jurisdictions) to pass a driving test without being taught. A fair number of candidates gained 1 mark by pointing out that the existence of driving tests has probably reduced the number of unsafe drivers on the road, even if it has not eliminated them.

A number of candidates wrongly identified the reference to 'incompetent drivers' as an instance of *ad hominem*. The passage's mention of 'poor people' was sometimes identified as an appeal to emotion, but such a judgement involved reading too much into the passage. A popular wrong answer was the generalisation in paragraph 4 that in certain countries 'everyone' accepts the need to bribe public officials, which does not weaken the reasoning; however, few candidates spotted the related flaw of moving from a claim about bribery in 'many countries' to a general recommendation about abolishing driving tests.

Some candidates understood the significance of the expression 'unstated assumption' and correctly identified one or more such assumptions which significantly weakened the reasoning; as on previous occasions, however, many misunderstood the phrase 'unstated assumption' to mean 'unsupported statement' and contradictorily claimed that quotations from the passage were unstated.

As on previous occasions, no credit was given for literary criticism, for comments about strengths or for criticism of the argument for being one-sided or lacking statistical support for its claims. No credit was given for saying that the author should have given examples of countries where bribery is widespread or for wrongly claiming that the author assumes all good test-takers are bad drivers and all good drivers are bad at taking tests.

- (d) Although the claim to be supported or challenged was slightly more complex than usual, the subject matter appeared to be familiar, and many candidates achieved 5 marks out of 5. It was harder to challenge this claim than to support it. A good number of candidates strengthened their arguments by making appropriate use of additional argument elements. A few candidates wrote their opinions, without turning them into an argument, thereby achieving 1 mark out of 5, while a few others argued for or against driving tests and were awarded 0 marks. Several candidates contradicted themselves by arguing that because driving is necessary in many situations we should not limit driving to those circumstances. Some candidates commented on the difficulty of defining 'necessary', which would be a major issue if one were contemplating legislation, but this question focused on personal choices, with no suggestion that they should be enforceable by law.

THINKING SKILLS

Paper 9694/22
Critical Thinking

General comments

There were many well-prepared able candidates achieving high marks. At the other end of the mark range, however, there was with quite a long 'tail' of candidates who were poorly prepared for the paper and who struggled to reach a mark in double-figures. Well-prepared candidates seemed to respond well to the issues raised by the questions and were able to tackle them effectively.

Key messages

- Some candidates identify the content of the source(s) that allows an inference to be made (which is often the answer to a question) but simply repeat it without making the inference. The effect of this is that they tend to repeat what the source says, leaving the significance still implicit and therefore fail to answer the question effectively.
- Candidates would be well-advised to always attempt **Question 3(d)** in the event of getting short of time.
- Some candidates are looking for answers from information in the sources where the question requires the application of their own thinking, for example **Question 2(a)**.

Comments on specific questions

Question 1

- (a) Only a minority of candidates tackled this question effectively. Some interpreted it as a credibility question. Others saw it as an *ad hominem* attack on Dr Kitty. Because the author was trying to establish with evidence that Dr Kitty was a con-man, this was not an argument with an *ad hominem* flaw. Those who saw it as effectively supporting the claim tended to just repeat what was in the passage and rely on accepting the implications the author intended. The key to answering the question was to see that if one questioned the implications, it was apparent that they were dubious, e.g. that an on-line course is necessarily suspect.
- (b) Most candidates gained at least one mark on this question. Candidates who gained 2 marks were more successful in making explicit how inheriting the money challenged Kitty's claim that his success was due to gaianomics.
- (c) This question was answered quite well with a significant number of candidates seeing the problem with a defence of gaianomics which seemed to give an account of a rejection of a dream rather than the realisation of it. Other points about possible bias and it being simply coincidental also featured frequently.
- (d) While most candidates argued that Kitty was a con-man, a significant minority argued the reverse. Some candidates got rather confused when they raised the admittedly difficult question of whether the book having a 'placebo' effect meant it was a con-trick or not. Some argued that Kitty knew that gaianomics was complete rubbish but that he used this as a device to write a self-help book which did help people. This meant he was not a con-man as such. Others disagreed and saw this as cynical and manipulative. Candidates who took this approach, irrespective of their final judgement, often made good use of inferential reasoning about how he may have used his psychological knowledge.

Question 2

- (a) This was quite an easy question with most candidates getting at least two reasons why governments would not tackle coastal erosion. Some candidates were helped by the fact that, on this occasion, relevant points could be gleaned from the sources. However, as noted in the key messages above, this is not normally an effective way of tackling questions like this.
- (b) Only a minority of candidates achieved 4 marks. The easiest way to achieve 4 marks was to identify and explain two factors that made global warming policy ineffective as a solution. Whilst the question did not explicitly ask one to do this, candidates need to look at the mark allocation to give them a sense of the length of answer required. A number of candidates seemed to think it was enough to say that coastal erosion was a natural phenomenon. This would not, in itself, explain why tackling global warming would be ineffective. Others raised issues of difficulties of implementation. However, the word 'achievement' should have implied successful implementation of the policy.
- (c) This question was done quite well with most candidates seeing that living on the coast did not necessarily mean living on a cliff subject to erosion and then going on to explain why this might not be the case.
- (d) Most candidates agreed with the proposition and there was plenty of evidence in the sources that building houses on cliff tops is unwise. However, some achieved a more nuanced conclusion by suggesting the claim was too sweeping and that there are probably some cliffs that do not erode because the rock is very hard. A small minority of candidates questioned the assumption that a cliff was necessarily by the sea. This was good critical thinking and was rewarded.

Question 3

- (a) Only a small minority of candidates successfully identified the conclusion and most achieved 0 marks. A rather greater number identified it as an intermediate conclusion and had therefore got part of the way to the correct answer. Such candidates would have benefited from testing whether this statement worked better as a reason to support a conclusion or a conclusion following from the reasoning. Applying this test should have led them to correctly identifying 'In the modern world one needs to get real and get planning' as the main conclusion.
- (b) In spite of not identifying the main conclusion, most candidates did identify two intermediate conclusions and seemed to therefore understand the major thrust of the argument.
- (c) By far the most identified flaw was the *ad hominem* in paragraph 2 but the assumptions in paragraphs 3 and 4 and the inconsistencies in paragraph 4 also figured frequently in the 4 or 5 mark answers. Very few candidates managed to explain the conflation in paragraph 1. Some weaker candidates challenged the statement by suggesting there was nothing wrong in setting off on a journey and not knowing where you were going. Candidates who achieved 0 marks tended to fall into two categories. Either they knew what they should be doing but could not do it – such answers tended to consist of many technical terms for flaws wrongly applied – or they did not understand the nature of the exercise. This second category of answer tended to challenge statements or comment on questions of style. Such answers are now in the minority.
- (d) Most candidates suggested clocks were necessary for life to function and that we would therefore be unhappy if we did not have them. Some thoughtful answers took the more difficult course of arguing for the proposition via the argument that 'clock time' is only necessary in a ruthless dehumanised technological society and therefore we would be happier if they had not been invented because it would mean we were in a different sort of society. On the other side, more sophisticated answers moved on from explaining things like how we'd never be able to catch a bus any more to suggesting time is an existential parameter and inventing clocks does not invent time or time pressure.

THINKING SKILLS

Paper 9694/23
Critical Thinking

General comments

There were a number of candidates who were poorly prepared for the paper and who struggled to reach a mark in double-figures. Well prepared candidates seemed to respond well to the issues raised by the questions and were able to tackle them effectively.

Key messages

- Some candidates identify the content of the source(s) that allows an inference to be made (which is often the answer to a question) but simply repeat it without making the inference. The effect of this is that they tend to repeat what the source says leaving the significance still implicit and therefore fail to answer the question effectively.
- Candidates would be well-advised to always attempt **Question 3(d)** in the event of getting short of time.

Comments on specific questions

Question 1

- (a) The most popular answers were that we could not be sure that the body was Diesel's and/or that the crew may have been lying for a variety of motives. A minority of candidates made the point that it meant no autopsy could be done to establish cause of death. Most candidates obtained at least 1 mark on this question.
- (b) Some less good responses suggested it actually strengthened the idea that the oil industry was responsible for his murder. Whilst one could eventually arrive at the conclusion that this evidence was not as significant as at first sight, one could not draw this conclusion immediately. Only a minority of candidates made the point that the potential use of oil to power diesel engines may not have been apparent at the time of Diesel's death.
- (c) Many candidates suggested that Source C gave grounds for believing Diesel was murdered rather than that he committed suicide, though some rather over-stated this. Good answers also reinforced this point by suggesting the information in Source C suggested Diesel's prospects were looking better, which may have deterred him from the idea of suicide.
- (d) Answers were fairly evenly split between those who thought he was murdered and those who thought he committed suicide. A number of candidates made the false inference that his clothes being laid out on the bed and him asking to be woken at 6 a.m. was conclusive evidence that he did not intend to commit suicide. Answers focused on the murder scenario tended to become speculative, e.g. he was thrown overboard by German agents.

Question 2

- (a) Not many candidates made the obvious point that if vaping was an alternative to smoking then one would expect a corresponding decline in smoking as vaping increased. Answers tended to offer explanations for why smoking had not decreased in spite of the increase in vaping. The two most common answers were that it may be true that vaping in young people encourages smoking, as suggested in Source B, and that vaping may be a totally different activity to cigarette smoking appealing to a totally different audience. Credit was given to these answers even though they did not quite focus on the question as originally asked.

- (b) Only a minority of candidates managed to focus on the two relevant pieces of evidence about brain damage in younger vapers, which would not apply to older vapers, and the dangers of overheated flavouring agents which could be avoided if certain steps were taken. Often only one piece of evidence was correctly identified
- (c) Many candidates were able to suggest two explanations for this increase, with increased advertising and fashionableness being popular answers.
- (d) Most candidates agreed with the proposition and the evidence in the sources did point to this conclusion. Even if the evidence that there was a risk was inconclusive, this also meant one could not definitely conclude that there was not any risk. A number of candidates identified the vested interest in Source A, which was the source which argued most strongly that vaping was a risk-free alternative to smoking.

Question 3

- (a) Many candidates got a 1 mark answer here, but only a minority managed to isolate the actual main conclusion.
- (b) The vast majority of candidates identified at least one intermediate conclusion and a significant number managed to identify three.
- (c) Whilst there were not many 5 mark answers, a reasonable number of candidates identified the appeal to history in paragraph 1 and the assumptions that work involves exercise / does not damage health in paragraph 3. Less successful answers were split between those who were unable to correctly identify assumptions and flaws, but did understand that this is what they were meant to be doing, and those who did not understand what this question requires. The latter were in a minority and typically commented on style or challenged what was being said.
- (d) A number of candidates misinterpreted this proposition as one about what science *could* achieve as opposed to what science *should* achieve. Such arguments often became rather lengthy digressions on the wonders of science. Of those who correctly identified the nature of the proposition, most argued against it, either pointing to the problems that an increasingly ageing population would bring, the poor quality of life of such elderly people or more important priorities for science, e.g. combating global warming. The minority who thought it should be a key goal tended to stress how much more people could achieve if they lived longer.

THINKING SKILLS

<p>Paper 9694/31 Problem Analysis and Solution</p>
--

Key messages

The pressures of the clock make time management a key element of success in this paper. Many candidates appeared to be casualties of this on **Question 4**, where the logic of the board game required great care rather than great insight. 'Great care' is difficult when you are running out of time. One sensible strategy is for candidates to allocate themselves less time for each question than the full 120 minutes divided proportionately (e.g. 20/20/30/30 for **Questions 1/2/3/4**) and stick to this. Then there is room at the end (the remaining 20 minutes) to apply concentrated effort to the questions which are predictable (and leave the ones which require insight, inferences, and imagination).

The ability to make exhaustive lists is a key problem solving skill that appears in many forms – in the problems here, and in any that depend upon consideration of all options. To be flexible and dependable in using it, it is worth drawing students' attention to its ubiquitous presence, and to enforce sensible organised methods for tracking which elements of a list have been considered and which have not.

General comments

On the whole candidates were able to engage with the contexts presented. Almost all candidates made reasonable attempts at some parts of every question.

None of the questions was particularly aided by a calculator – and this may have enabled the less numerate candidates to avoid some of the spectacular 'rounding' errors that have occurred in previous series. There was the usual balance between numerical answers and explanation – and candidates offered reasonably clear working on the whole, which did enable them to pick up partial credit if they made minor arithmetic errors.

Comments on specific questions

Question 1

This question applied two different problems to a similar context: the first part involved the selection and arrangement of certain quantities to satisfy a rule; the second part (independent of the first) involved careful application of a fairly complex rule to optimise the outcome.

- (a) This question was an easy introduction to the context, and almost all candidates answered it correctly.
- (b) This question could be tackled in exactly the same way as **part (a)**, or approached by consideration of the changes that had been made (+0.4MB, +0.3MB and –0.3MB). Those candidates that used a 'first fit' approach for both parts (starting with the largest) found themselves needing to start afresh in **(b)**. Many different approaches were possible – and any that achieved an appropriate separation were awarded the mark.
- (c) This question offered a minor extension to the situation, asking candidates to optimise the distribution of a new set of sizes. Success depended on pairing the 6.4 MB and 2.7 MB files – and most candidates did this successfully.

- (d) The final part of the question did not depend on the previous parts – it took a similar context, and required candidates to apply an interdependent group of rules (regarding capacities, proportional rates, fixed rates and additional costs) to find a minimum cost for 350 envelopes. A gratifying number of candidates showed their working in a transparent way, and gained partial credit for their efforts where they did not manage to deliver a correct final answer. Very few managed to apply all the rules without error and gain 5 marks. The rule that seemed to cause the most problems was the additional postage cost, which needed to be understood alongside the fixed postage cost.

Question 2

This question involved looking at a taxi driver's profits as each of the different parameters in the relationship changed. The relationship between the average journey distance, average speed, fuel price and charge per kilometre was not hard to grasp – but combination of different rates (km per hour, \$ per km) needed sensitive analysis.

- (a) This question was designed to confirm the structure of the problem to candidates. To gain the mark candidates did need to show each of the separate components that combined to make \$8. A number of answers offered little more than $2 + 10 - 4 = 8$ which was not awarded any marks, since there was no evidence that the candidate knew where the \$4 came from.
- (b) This question reversed the logic of part (a), asking for a distance which would yield a given profit. The most common error was to forget about the fuel costs of the journey back – giving the answer \$15. Almost all those who managed to answer (a) correctly also managed (b).
- (c) This question gently introduced speed as a parameter, using the profit from (a) and encouraging candidates to find the number of journeys, as a stepping stone to the profit. Those who forgot about the fuel costs on the journey home (giving \$384) obtained no marks here, but were awarded a mark in (d) if they managed everything else right.
- (d) This question could be tackled with 'hard data' (calculating the profit for a day of 20 km journeys, and comparing) or with careful analysis of how the profit was affected by the fixed fare. The latter involved less effort, but depended upon careful expression. The former was more popular.
- (e) This reversed the logic of (c) and (d), and required something approaching an algebraic formulation of the relationship, or a breakdown of the profit into individual journeys. Most candidates managed to offer a correct answer (2 marks), or left no comprehensible working at all (0 marks).
- (f) With the alteration of one of the other rates in the question, candidates had to analyse how the profit was related to the charge per kilometre. This amendment of the underlying model put a few candidates off attempting it. Half of those who did attempt it (by intelligent guesswork and refinement of their solution, or by inference) answered correctly.

Question 3

This problem involved a daily-altering Caesar cipher, and the inferences one could make from it. Candidates were not expected to be familiar with such codes – and the vast majority found it easy to engage with.

- (a) (i) The opening question involved a careful selection of the cipher text for the word CIPHER – some candidates applied it the wrong way round (KQVPMH), which was a harder task, and showed little appreciation of the example given directly above the question.
- (ii) The correct placing of the letters A, P, R, I and L on the 12th, 13th, 14th, 15th and 16th letters, and the placement of the remaining letters in alphabetical order, starting with B on the 17th gave a new code. Consistent 'reverse application' of the cipher was awarded 1 mark here.
- (iii) This was a substantially harder question, involving careful inferences from the desired cipher text (R, O and S on their own letters) to the months without those letters (MAY, JUNE, JULY), and then the dates on which O will be the 21st letter. Candidates found this one of the hardest questions on the paper.

- (b) This was the first question which required candidates to think about what happened on dates after the 26th of any month. Because the cipher repeats at this point, there are only $(26 \times 12 =)$ 312 different ciphers, not 365.
- (c) (i) This question involved reverse engineering the cipher to start at A, which could be done fairly easily by counting back 6 letters from 26 (O, C, T, B, E and R).
- (ii) This was a substantially harder question, since it involved correctly deciding what to do with all the months with an A in (JANUARY, FEBRUARY, MARCH, APRIL, MAY and AUGUST), and what do with the dates after 26th in any month. Partial credit was given generously – for any clear evidence that the candidates appreciated that there would be 12 dates, before the months with an A were considered.
- (d) A pair of letters and an explanation were needed here. Quite a few candidates did not manage to clearly explain the need for the letters to be consecutive and to not appear in the letters of any month.
- (e) Many candidates tackled this clearly and effectively, often beginning with a partially filled grid of the letters that could be identified from the message. This was awarded a mark even if the necessary inferences from this were not carried out. The question was tackled effectively by nearly half the candidates.

Question 4

This question depicted a board game, and required candidates to track the choices of the players, and consider their optimal moves. As is intended with complex data questions, the initial task was to interpret the page of instructions and examples in order to apply them correctly. Once that had been accomplished, the logical space in which the play occurred was fairly limited, and many candidates gave sensible answers to most of the questions – however, many appeared to have tackled this question last, and there was evidence that the lack of time hampered candidates' approaches to the final parts.

- (a) The opening question invited candidates to place the given words around the central column of E, I, N, O and T. The word FOUR was fixed, but there were quite a few alternatives for the remaining four words.
- (b) This question required a careful listing of all the places which would enable an N to spell a number word. Candidates found this quite hard, with many giving the places used in SEVEN, NINE and TEN, but not ONE.
- (c) This question began like part (a), fitting the words onto the grid (a unique solution this time), and then required a careful allocation of points. Tabulation of the answers made this fairly easy.
- (d) This was the first question which involved rigorous tracking of which tiles had been used, and which remained. Independently, candidates needed to ascertain which tiles could be used, and then select the three that couldn't. This latter selection process required careful consideration of each of the number words, and many made minor errors (perhaps resulting from time pressures).
- (e) (i) Careful tracking of the possible play revealed the second S was not playable. Some candidates appeared to be guessing.
- (ii) The most popular wrong answer was 'There were not enough of the right tiles in the bag' – which was too vague to be creditworthy.
- (iii) Those who managed (ii) tended to complete this correctly as well.
- (f) This question was the one candidates scored least well on, on the entire paper – but it seems more likely that this was a result of time restrictions than it requiring particularly tough logic. The inferences needed here did not involve any great leaps, or awkward reverse engineering, and a careful exhaustive explanation was fairly easy to deliver. About a third of those who tackled this were successful.

THINKING SKILLS

Paper 9694/32
Problem Analysis and Solution

Key messages

- The information in the text of questions is almost always needed at some point; most candidates would benefit from checking that they have found some use for all of the information and have not neglected important detail.
- Some candidates would gain more marks by labelling the numbers in their working, thereby allowing examiners to see evidence for partial credit more often. Doing so should also help them to detect errors.
- Candidates should attempt the parts in the order in which they are given, as earlier parts often highlight detail that is essential for a fully correct solution. However, sometimes in a scenario an avenue is explored and then excluded from further consideration; careful reading of the question is necessary to avoid carrying over assumptions or conditions that are not valid for the current part question.
- 'Explain' needs some reasoning, usually with some numeric justification, and not simply rephrasing the statement given.
- When the answer is given in the question, examiners are generally looking for a demonstration of how it is obtained; so there is no point in highlighting it, and every reason not to cross-out the working.

Comments on specific questions

Question 1

This question asked for the maximum or minimum of participants or medals for a squad with different teams, as well as finding a suitable arrangement of players into teams for various events. It was important for candidates to take care over the terminology and not to use 'team' as a generic word. Most candidates found this question harder than expected; many opted to start with a different one.

- (a) This asked for the least number so that a full complement of teams could be fielded, (so for each event multiplying the number of team by members and then taking the maximum). Many candidates simply took the minimum for an event.
- (b)(i) The number of cyclists in a particular team could be deduced by considering the total needed and how they were arranged either singly or doubling up. Few candidates did this correctly.
- (ii) An arrangement satisfying various constraints was asked for, using letters to label cyclists. Few candidates offered arrangements that satisfied 'all other cyclists must take part in exactly two events', highlighting the lack of attention to detail in the previous part. Some listed the events for each cyclist rather than the cyclist for each event, losing the benefit of the suggested labelling and requiring considerably more writing.
- (c) This sought the greatest possible medal total, which required noticing that only three teams in any event get medals. Almost all candidates responded, but only a third gave the correct answer, despite this deduction being straightforward. Better responses took account of this constraint.
- (d)(i) The rest of the question considered the cases where we are told there are 7 gold medals but not what they were for. The first part asks for three ways in which this can happen. A quarter of candidates offered no response, and a few only offered one or two.

- (ii) The calculation of the maximum number of silver medals for the squad required noticing that those who gained gold would not get silver. A third of candidates offered no response, and many of the responses simply assumed that if they had 7 gold then 7 was as many medals as could be obtained.
- (e) This sought the lowest maximum of the possible cases, and called for careful counting. A third of candidates offered no response.

Question 2

Different costs in two countries as well as a varying exchange rate can require an allowance to supplement salaries. Different ways to calculate an allowance are explored, along with the consequences of changes in the good use for comparison. Better responses made clear what differences or ratios were being discussed, and which currency was considered.

- (a) A total price for a basket of goods using the cheapest source for each was called for. This involved taking the smallest number on each row from the relevant three columns.
- (b) This looked at average prices, and called for the median price of each item to be determined and the total of all items in the basket calculated. Few candidates used the median, and many of those either took the median of the totals or the median of all the prices rather than the median for each item (all of which are the same for the mean).
- (c) This looked at the impact of 'loss-leaders' (although that term was not used). Some candidates used a laborious 'try everything' approach instead of calculating the cost for each supermarket and then identifying the single case that matched the difference.
- (d) This considered the impact of changes under different scenarios on how the allowance could be calculated: as an absolute difference or as a ratio. Twice as many candidates did not attempt the ratio than did not respond to the difference. A few worked out the correct figures but then, with lack of words identifying the numbers, applied them the wrong way around.
- (e) The direction of change when one parameter was decreased was asked for; few responses gave a correct direct answer.
- (f) This explored the impact on the spending power on staff when the exchange rate changed so that no allowance was needed. No calculations were needed, but there was no response from a quarter of candidates.

Question 3

Given constraints on room size, times and costs, this question looked at options for parties. The lowest rate for a given model giving a specified profit was wanted. Candidates needed to be clear what was a cost for the organiser and what was income (and cost to someone else).

- (a) The minimum price for a given party for a specified time and number of guests involved determining the cheaper of the options that could accommodate the number of guests. All candidates answered and almost all gave the correct answer.
- (b) This involved taking the difference between Alice's income and her costs in a specific case and correctly identifying whether it was a profit or a loss. Better responses included all the costs. A few just gave a number and did not make it clear whether it was a profit or a loss.
- (c) The minimum number of guests to make a profit was asked for. Many candidates gave the profit (in dollars) rather than the number of guests. Some looked only at the minimum numbers for the rooms where these were specified and ignored the case where there was no minimum.
- (d)(i) Many candidates ignored the constraint that parties were between 2 and 5 hours. Nobody guessed the wrong end.
 - (ii) Some candidates used algebra, but made difficulty for themselves by using $(\text{income} - \text{cost})/\text{cost} > 20\%$ rather than a simpler formulation.

- (e) This part wanted what was identified as a reduced rate. Candidates who offered an *increased* rate could have checked this. Some sought to reduce the profit by 5% rather than add 5% to costs. It was necessary to consider the 30 guests at full price as distinct from the 20 guests at a reduced price, and few made the distinction. A quarter offered no response.

Question 4

A competition with a curious scoring method was presented, along with incomplete tables of scores. Since scores with halves and quarters were involved, premature rounding caused errors.

- (a) The theoretical maximum score that could be obtained in a round was required. Some candidates would have benefited from noticing that some of the actual scores were greater than their suggested maximum.
- (b)(i) Full marks required correctly determining the original total, adding 2×3 , and applying the penalty for contestants making a second attempt. Some candidates noticed that their responses involved more than the maximum but did not deduce that they had made a mistake.
- (ii) The score from the first attempt was easier to calculate as it had no addition and no penalty; most candidates calculated this correctly.
- (c)(i) This involved finding the three cases which were multiples of 2.5. Most responses included the two obvious cases (with 0.5).
- (ii) A second attempt by a contestant would involve a 5-point penalty. One of the scores was not a multiple of 3 nor of 2.5. Most responses were based on this, but some correctly observed that the rate was above 2.5 in the relevant case, and so must be 3 with penalty.
- (d) A third of candidates offered no response to this part, and many candidates addressed a different and easier question: why the individual judge's scores could not be determined from the total. It was sufficient to note that both 2.5 and 3 are factors of 45. Some candidates wasted time exploring other divisors without taking account of the limit of $3 \times 8 = 24$ on the multiplier. Many of those candidates who scored well on other questions did not gain this mark, possibly attempting a short-cut (but leaving no evidence of their reasoning).
- (e) The grand total involved more than adding three scores for tricks in a table; half of an earlier score was needed, but only about half of the candidates took this into account.
- (f)(i) The instruction was 'show that at least 2....' so for both marks it was necessary to show that 1 was not sufficient. Many responses only offered how it could be done with three, and a few used more. One third of candidates did not attempt this part.
- (ii) This involved finding the highest score that had been or could still be obtained by other contestants, not just the one mentioned in (f)(i), and comparing with Cuthbert's. Better responses looked at the general case and not the specific one explored in the previous part.

THINKING SKILLS

<p>Paper 9694/33 Problem Analysis and Solution</p>
--

Key messages

- The information in the text of questions is almost always needed at some point; most candidates would benefit from checking that they have found some use for all of the information and have not neglected important detail.
- Some candidates would gain more marks by labelling the numbers in their working, thereby allowing examiners to see evidence for partial credit more often. Doing so should also help them to detect errors.
- Candidates should attempt the parts in the order in which they are given, as earlier parts often highlight detail that is essential for a fully correct solution. However, sometimes in a scenario an avenue is explored and then excluded from further consideration; careful reading of the question is necessary to avoid carrying over assumptions or conditions that are not valid for the current part question.
- 'Explain' needs some reasoning, usually with some numeric justification, and not simply rephrasing the statement given.
- When the answer is given in the question, examiners are generally looking for a demonstration of how it is obtained; so there is no point in highlighting it, and every reason not to cross-out the working.

Comments on specific questions

Question 1

This question asked for the maximum or minimum of participants or medals for a squad with different teams, as well as finding a suitable arrangement of players into teams for various events. It was important for candidates to take care over the terminology and not to use 'team' as a generic word. Most candidates found this question harder than expected; many opted to start with a different one.

- (a) This asked for the least number so that a full complement of teams could be fielded, (so for each event multiplying the number of team by members and then taking the maximum). Many candidates simply took the minimum for an event.
- (b)(i) The number of cyclists in a particular team could be deduced by considering the total needed and how they were arranged either singly or doubling up. Few candidates did this correctly.
- (ii) An arrangement satisfying various constraints was asked for, using letters to label cyclists. Few candidates offered arrangements that satisfied 'all other cyclists must take part in exactly two events', highlighting the lack of attention to detail in the previous part. Some listed the events for each cyclist rather than the cyclist for each event, losing the benefit of the suggested labelling and requiring considerably more writing.
- (c) This sought the greatest possible medal total, which required noticing that only three teams in any event get medals. Almost all candidates responded, but only a third gave the correct answer, despite this deduction being straightforward. Better responses took account of this constraint.
- (d)(i) The rest of the question considered the cases where we are told there are 7 gold medals but not what they were for. The first part asks for three ways in which this can happen. A quarter of candidates offered no response, and a few only offered one or two.

- (ii) The calculation of the maximum number of silver medals for the squad required noticing that those who gained gold would not get silver. A third of candidates offered no response, and many of the responses simply assumed that if they had 7 gold then 7 was as many medals as could be obtained.
- (e) This sought the lowest maximum of the possible cases, and called for careful counting. A third of candidates offered no response.

Question 2

Different costs in two countries as well as a varying exchange rate can require an allowance to supplement salaries. Different ways to calculate an allowance are explored, along with the consequences of changes in the good use for comparison. Better responses made clear what differences or ratios were being discussed, and which currency was considered.

- (a) A total price for a basket of goods using the cheapest source for each was called for. This involved taking the smallest number on each row from the relevant three columns.
- (b) This looked at average prices, and called for the median price of each item to be determined and the total of all items in the basket calculated. Few candidates used the median, and many of those either took the median of the totals or the median of all the prices rather than the median for each item (all of which are the same for the mean).
- (c) This looked at the impact of 'loss-leaders' (although that term was not used). Some candidates used a laborious 'try everything' approach instead of calculating the cost for each supermarket and then identifying the single case that matched the difference.
- (d) This considered the impact of changes under different scenarios on how the allowance could be calculated: as an absolute difference or as a ratio. Twice as many candidates did not attempt the ratio than did not respond to the difference. A few worked out the correct figures but then, with lack of words identifying the numbers, applied them the wrong way around.
- (e) The direction of change when one parameter was decreased was asked for; few responses gave a correct direct answer.
- (f) This explored the impact on the spending power on staff when the exchange rate changed so that no allowance was needed. No calculations were needed, but there was no response from a quarter of candidates.

Question 3

Given constraints on room size, times and costs, this question looked at options for parties. The lowest rate for a given model giving a specified profit was wanted. Candidates needed to be clear what was a cost for the organiser and what was income (and cost to someone else).

- (a) The minimum price for a given party for a specified time and number of guests involved determining the cheaper of the options that could accommodate the number of guests. All candidates answered and almost all gave the correct answer.
- (b) This involved taking the difference between Alice's income and her costs in a specific case and correctly identifying whether it was a profit or a loss. Better responses included all the costs. A few just gave a number and did not make it clear whether it was a profit or a loss.
- (c) The minimum number of guests to make a profit was asked for. Many candidates gave the profit (in dollars) rather than the number of guests. Some looked only at the minimum numbers for the rooms where these were specified and ignored the case where there was no minimum.
- (d)(i) Many candidates ignored the constraint that parties were between 2 and 5 hours. Nobody guessed the wrong end.
 - (ii) Some candidates used algebra, but made difficulty for themselves by using $(\text{income} - \text{cost})/\text{cost} > 20\%$ rather than a simpler formulation.

- (e) This part wanted what was identified as a reduced rate. Candidates who offered an *increased* rate could have checked this. Some sought to reduce the profit by 5% rather than add 5% to costs. It was necessary to consider the 30 guests at full price as distinct from the 20 guests at a reduced price, and few made the distinction. A quarter offered no response.

Question 4

A competition with a curious scoring method was presented, along with incomplete tables of scores. Since scores with halves and quarters were involved, premature rounding caused errors.

- (a) The theoretical maximum score that could be obtained in a round was required. Some candidates would have benefited from noticing that some of the actual scores were greater than their suggested maximum.
- (b)(i) Full marks required correctly determining the original total, adding 2×3 , and applying the penalty for contestants making a second attempt. Some candidates noticed that their responses involved more than the maximum but did not deduce that they had made a mistake.
- (ii) The score from the first attempt was easier to calculate as it had no addition and no penalty; most candidates calculated this correctly.
- (c)(i) This involved finding the three cases which were multiples of 2.5. Most responses included the two obvious cases (with 0.5).
- (ii) A second attempt by a contestant would involve a 5-point penalty. One of the scores was not a multiple of 3 nor of 2.5. Most responses were based on this, but some correctly observed that the rate was above 2.5 in the relevant case, and so must be 3 with penalty.
- (d) A third of candidates offered no response to this part, and many candidates addressed a different and easier question: why the individual judge's scores could not be determined from the total. It was sufficient to note that both 2.5 and 3 are factors of 45. Some candidates wasted time exploring other divisors without taking account of the limit of $3 \times 8 = 24$ on the multiplier. Many of those candidates who scored well on other questions did not gain this mark, possibly attempting a short-cut (but leaving no evidence of their reasoning).
- (e) The grand total involved more than adding three scores for tricks in a table; half of an earlier score was needed, but only about half of the candidates took this into account.
- (f)(i) The instruction was 'show that at least 2....' so for both marks it was necessary to show that 1 was not sufficient. Many responses only offered how it could be done with three, and a few used more. One third of candidates did not attempt this part.
- (ii) This involved finding the highest score that had been or could still be obtained by other contestants, not just the one mentioned in (f)(i), and comparing with Cuthbert's. Better responses looked at the general case and not the specific one explored in the previous part.

THINKING SKILLS

<p>Paper 9694/41 Applied Reasoning</p>
--

Key messages

- In **Question 1** many candidates wrote longer answers than needed.
- In **Question 2** many candidates were paraphrasing the argument elements or evaluating the argument. Marks are awarded only for precise identification of argument elements.
- In **Question 3** many candidates are still attempting the wrong task. They are asked to evaluate the reasoning, not to argue against it or to write their own opinions on the topic. Study of previous mark schemes should reveal the kinds of answers that are credited.
- In **Question 4**, candidates can only achieve the highest marks if they engage critically with the documents provided.

General comments

It was evident that a large number of candidates had had very little training in the rudiments of critical thinking or the format of the Applied Reasoning paper. This affected their ability to score marks in the shorter questions because of a lack of vocabulary associated with reasoning or understanding of the requirements of the questions.

Most candidates appeared to have enough time to finish the paper. Those who appeared to be rushing on **Question 4** had often spent a long time on other questions, achieving very few marks, where marks would have been achieved more easily on **Question 4**. In previous reports it has been noted that more candidates were writing answers whose length reflected the mark allocation more closely. Unfortunately, this trend has not continued this session and many candidates were writing responses to **Questions 1, 3** and, to a lesser extent, **2** that were too long.

The standard of candidates varied but there was evidence that many candidates had not been well prepared. Many did not know what they were being asked to do, particularly in **Questions 2** and **3**.

Comments on specific questions

Question 1

Many candidates attempted to criticise the data or the inferences drawn from it. Most achieved 1 or 2 marks but only a few achieved more. The most commonly seen creditworthy response noticed that feeling able to run is not the same as physical fitness. All other points on the mark scheme were seen but most were not credited very often. Some candidates noted the two y-axes but expressed their concerns in terms of having difficulty understanding the graph, rather than the inappropriateness of a point of intersection. It was common for candidates to regard the fact that the graph did not show each volunteer's data individually as a problem. As ever, some responses discussed the lack of a source or sources given and achieved no credit.

Question 2

This question rewarded the well-prepared candidate. Those who knew what was expected and attempted an analysis of the argument usually gained at least half the marks, and often more, easily. Many candidates struggled to unpick elements from within a single sentence. It was common, even for moderately well-prepared candidates, to quote a CA and the MC and label the whole response as MC. Likewise, many identified the first IC along with the reason that supports it as an IC. Words like 'because' or 'however' are unlikely to appear in the middle of an argument element. Many candidates provided a non-creditworthy summary or gist. Some seemed unaware that quoting from the text is an appropriate, indeed a required, way

to answer this question. A small minority attempted to evaluate the reasoning, which they were invited to do in **Question 3**.

Question 3

Full evaluation of this passage was quite challenging. The better prepared candidates attempted to evaluate the passage, but many are still listing a series of counter-arguments to points in the passage. Most candidates who did attempt to apply their evaluation skills gained 2 marks for identifying the slippery slope in paragraph 3 but rarely gained more. A few, usually high-scoring candidates, understood what an assumption was and were able to identify several. Most candidates, however, appeared to regard any unsubstantiated claim as an assumption. It is worth reiterating that unstated assumptions cannot have been stated in the document. Many candidates offered counters to the claims in the passage; this was not credited.

Question 4

Candidates were required to use the documents and their own ideas to construct a reasoned case to support or challenge the conclusion that government decisions should represent the will of the people. Most candidates did focus their entire argument on supporting or challenging a conclusion – the majority being in support. However, many candidates struggled to reconcile ‘the will of the people’ with various interpretations of democracy and many argued for a conclusion that was about the benefits of democracy rather than the will of the people.

The majority of candidates scored below 11 marks because their reasons rarely went beyond what was written in the documents and the information given in the documents was used uncritically. Evaluation of documents was less common than in previous sessions but there were more attempts to consider counter positions and attempt to deal with them. A minority of candidates were able to produce a thoughtful and reasoned case that achieved higher marks. It was pleasing to see some candidates attempt to structure their arguments using strands of reasoning and intermediate conclusions and only a very few failed to state any conclusion at all. A small number of candidates simply described the contribution made by each document to the debate and achieved very few marks. As stated in previous reports, what is likely to get high marks is a persuasive argument with a clear structure that is supported by thoughtful, particularly critical, use of the documents and that thoughtfully considers and refutes relevant alternative viewpoints.

THINKING SKILLS

<p>Paper 9694/42 Applied Reasoning</p>
--

Key messages

- In **Question 1** many candidates wrote longer answers than needed.
- In **Question 2** many candidates are paraphrasing the argument elements or evaluating the argument. Marks are awarded only for precise identification of argument elements.
- In **Question 3** many candidates are still attempting the wrong task. They are asked to evaluate the reasoning, not to argue against it or to write their own opinions on the topic. Study of previous mark schemes should reveal the kinds of answers that are credited.
- In **Question 4**, candidates can only achieve the highest marks if they engage critically with the documents provided.

General comments

Most candidates appeared to have enough time to finish the paper. Those who appeared to be rushing on **Question 4** had often spent a long time on **Questions 1** or **3**, achieving very few marks, where marks would have been achieved more easily on **Question 4**. In previous reports it has been noted that more candidates were writing answers whose length reflected the mark allocation more closely. Unfortunately, this trend has not continued this session and many candidates were writing responses to **Questions 1, 3** and, to a lesser extent, **2** that were too long.

The standard of candidates varied but there was evidence that many candidates had not been well prepared. Many did not know what they were being asked to do, particularly in **Questions 2** and **3**.

Comments on specific questions

Question 1

Almost all candidates attempted to criticise the data or the inferences drawn from it. Most achieved 1 or 2 marks but some achieved more. The most commonly seen creditworthy response was to mention that SSRIs are used to treat conditions other than depression. Most other points on the mark scheme were seen and credited but less frequently. The only point examiners did not see was discussion about the ambiguity of the term 'sales'. The passage referred to a large-scale study so responses that questioned the sample size were not credited. Some candidates misinterpreted the passage and inferred that the increase in depression was within the group of celebrities only. As ever, some responses discussed the lack of a source or sources given and achieved no credit.

Question 2

This question rewarded the well-prepared candidate. Those who knew what was expected and attempted an analysis of the argument usually gained at least half the marks, and often more, easily. However, many candidates provided a non-creditworthy summary or gist. Some seemed unaware that quoting from the text is an appropriate, indeed a required, way to answer this question. A small minority attempted to evaluate the reasoning, which they were invited to do in **Question 3**.

Question 3

The better prepared candidates attempted to evaluate the passage, but many are still listing a series of counter-arguments to points in the passage. Most candidates who did attempt to apply their evaluation skills struggled to gain more than 2 marks. The most frequently credited weaknesses identified were weak analogy

between tobacco company claims and GMO manufacturer claims and, to a lesser extent, the appeal to ignorance in paragraph 4 and the *non-sequitur* in paragraph 5. A few, usually high-scoring, candidates understood what an assumption was and were able to identify several. Most candidates, however, appeared to regard any unsubstantiated claim as an assumption. Many candidates offered counters to the claims in the passage. This was not credited.

Question 4

Candidates were required to use the documents and their own ideas to construct a reasoned case to support or challenge the conclusion that we should encourage the development of GM foods. Most candidates did focus their entire argument on supporting or challenging the conclusion – the majority being in support. Fewer candidates than in previous sessions considered both sides equally and then decided on a conclusion at the last minute. In responses that adopted this latter approach it was difficult to award higher level marks in both the structure and quality skill sections.

The majority of candidates scored below 12 marks because their reasons rarely went beyond what was written in the documents and the information given in the documents was used uncritically. A minority of candidates were able to produce a thoughtful and reasoned case that achieved higher marks. It was pleasing to see some candidates attempt to structure their arguments using strands of reasoning and intermediate conclusions and only a very few failed to state any conclusion at all. A small number of candidates simply described the contribution made by each document to the debate. As stated in previous reports, what is likely to get high marks is a persuasive argument with a clear structure that is supported by thoughtful, particularly critical, use of the documents and that thoughtfully considers and refutes relevant alternative viewpoints.

THINKING SKILLS

Paper 9694/43
Applied Reasoning

Key messages

- In **Question 1** many candidates wrote longer answers than needed.
- In **Question 2** many candidates are paraphrasing the argument elements or evaluating the argument. Marks are awarded only for precise identification of argument elements.
- In **Question 3** many candidates are still attempting the wrong task. They are asked to evaluate the reasoning, not to argue against it or to write their own opinions on the topic. Study of previous mark schemes should reveal the kinds of answers that are credited.
- In **Question 4**, candidates can only achieve the highest marks if they engage critically with the documents provided.

General comments

Most candidates appeared to have enough time to finish the paper. Those who appeared to be rushing on **Question 4** had often spent a long time on **Questions 1** or **3**, achieving very few marks, where marks would have been achieved more easily on **Question 4**. In previous reports it has been noted that more candidates were writing answers whose length reflected the mark allocation more closely. Unfortunately, this trend has not continued this session and many candidates were writing responses to **Questions 1, 3** and, to a lesser extent, **2** that were too long.

The standard of candidates varied but there was evidence that many candidates had not been well prepared. Many did not know what they were being asked to do, particularly in **Questions 2** and **3**.

Comments on specific questions

Question 1

Almost all candidates attempted to criticise the data or the inferences drawn from it. Most achieved 1 or 2 marks but some achieved more. The most commonly seen creditworthy response was to mention that SSRIs are used to treat conditions other than depression. Most other points on the mark scheme were seen and credited but less frequently. The only point examiners did not see was discussion about the ambiguity of the term 'sales'. The passage referred to a large-scale study so responses that questioned the sample size were not credited. Some candidates misinterpreted the passage and inferred that the increase in depression was within the group of celebrities only. As ever, some responses discussed the lack of a source or sources given and achieved no credit.

Question 2

This question rewarded the well-prepared candidate. Those who knew what was expected and attempted an analysis of the argument usually gained at least half the marks, and often more, easily. However, many candidates provided a non-creditworthy summary or gist. Some seemed unaware that quoting from the text is an appropriate, indeed a required, way to answer this question. A small minority attempted to evaluate the reasoning, which they were invited to do in **Question 3**.

Question 3

The better prepared candidates attempted to evaluate the passage, but many are still listing a series of counter-arguments to points in the passage. Most candidates who did attempt to apply their evaluation skills struggled to gain more than 2 marks. The most frequently credited weaknesses identified were weak analogy

between tobacco company claims and GMO manufacturer claims and, to a lesser extent, the appeal to ignorance in paragraph 4 and the *non-sequitur* in paragraph 5. A few, usually high-scoring, candidates understood what an assumption was and were able to identify several. Most candidates, however, appeared to regard any unsubstantiated claim as an assumption. Many candidates offered counters to the claims in the passage. This was not credited.

Question 4

Candidates were required to use the documents and their own ideas to construct a reasoned case to support or challenge the conclusion that we should encourage the development of GM foods. Most candidates did focus their entire argument on supporting or challenging the conclusion – the majority being in support. Fewer candidates than in previous sessions considered both sides equally and then decided on a conclusion at the last minute. In responses that adopted this latter approach it was difficult to award higher level marks in both the structure and quality skill sections.

The majority of candidates scored below 12 marks because their reasons rarely went beyond what was written in the documents and the information given in the documents was used uncritically. A minority of candidates were able to produce a thoughtful and reasoned case that achieved higher marks. It was pleasing to see some candidates attempt to structure their arguments using strands of reasoning and intermediate conclusions and only a very few failed to state any conclusion at all. A small number of candidates simply described the contribution made by each document to the debate. As stated in previous reports, what is likely to get high marks is a persuasive argument with a clear structure that is supported by thoughtful, particularly critical, use of the documents and that thoughtfully considers and refutes relevant alternative viewpoints.