THINKING SKILLS

Paper 9694/21
Critical Thinking

Key messages

Candidates need to recognise what kind of answer is implied by each type of question and to answer accordingly. The best way of developing this skill is by studying previous question papers, mark schemes and Principal Examiner's reports.

In **Questions 1a, 1b, 1c, 2a, 2b** and **2c**, there are usually (but not invariably) two sides to the issue. In such cases, the mark is usually capped if candidates recognise only one side. They should be advised to look for two sides, but to be aware that in a few cases the correct answer may be "very useful", "not useful at all", or the equivalent.

It is almost always a good strategy to devote some time to thinking before beginning to write.

General comments

There was little or no evidence that any candidates had declined to engage with the examination in general. The lowest marks were awarded because of unclear thinking or claims being made without support.

There were very few "No Responses" and no instances of candidates apparently not having time to finish the paper.

As in previous sessions, strengths and weaknesses tended to be typical of particular Centres.

Comments on specific questions

Question 1

Unusually, this question referred to an actual historical debate, namely the fate of 'the princes in the Tower'. However, all the information required to answer the questions was provided in the sources. A few candidates were confused, and made suggestions such as that Sir John Tyrell was trying to protect Richard III (who was dead), that Richard III and Henry VII killed the princes together, or that Henry VII would not want to kill the princes because he was their father.

(a) Many candidates scored 2 marks out of 3 by noting that Tyrell's evidence was unreliable because he would be likely to say whatever his torturers wanted him to say. A significant minority claimed that the evidence was reliable, because under torture people always tell the truth; this answer was not credited. Some candidates, but not very many, achieved the third mark, by pointing out that Tyrell had excellent ability to know what had happened.

(b) Most candidates gave very little credence to the Richard III Society, on the basis that it was biased and/or lacked expertise. Relatively few gained the third mark, by pointing out that it is an academic society, which engages in serious research, or that their views were useful because they provided an alternative perspective on the history. Quite a lot of candidates unreasonably claimed that the society's views were not based on evidence, and this was not credited.

(c) There were several valid comments which could be made in response to this question, and most candidates spotted at least one of them. Some over-stated the significance, by claiming that the discovery of the skeletons "proved" that the princes had been murdered.

(d) As usual, in order to achieve Level 3, candidates needed both to evaluate the sources and to consider a possible alternative scenario. Slightly more candidates did this than in previous sessions. On this occasion, the sources themselves clearly hinted at several alternative scenarios, and most candidates discussed, or at least mentioned, one or more of them. Unsurprisingly, the most popular answers were that the princes were killed by Richard III or by Henry VII. Some candidates argued that the princes were murdered, without considering who might have been the murderer.

Question 2

This question was based on a recent genuine research project. The statistics in Source C and the information in Sources B and D were all taken from the project report. The information about physiognomy in Source A was also genuine.

(a) Most candidates achieved 2 marks by showing that the statistics in Source C did support the claim in Source D. There were several ways of achieving a third mark, and some candidates did so. A few candidates correctly and perceptively stated that the statistics strictly supported only a modified version of the claim, namely that unattractive people were more likely to *admit* to having committed crimes.

(b) Only a minority of candidates correctly identified the key point, namely that the claim was not a valid criticism of the research exercise, because allowance had been made for factors which might indicate poverty. However, there were several other valid comments which scored 1 mark, and quite a lot of candidates identified at least one of them.

(c) Not many candidates identified the key point, namely that comments about arrests and convictions were not directly relevant, because the research data were self-reported. However, there were several other valid points that could be made, and a fair number of candidates achieved at least 1 mark.

(d) Most candidates did make some reference to the source material, and a good many of them made some relevant criticisms. Very few discussed the transition from the claim of the research (that unattractive people are more likely to commit crime than attractive people) to the claim in the question (that it is possible to recognise potential criminals by their faces).

Question 3

In general, this argument made some reasonable points, but most candidates considered the conclusion to be over-stated.

(a) Most candidates, although by no means all, correctly identified the main conclusion of the argument. Although it was fairly prominent, it is unlikely that anyone would have guessed this as the right answer, rather than identifying it by analysing the argument correctly. A few candidates were awarded 1 mark for paraphrasing the conclusion or incorporating additional material.

(b) As usual, the correct answers to this question were intermediate conclusions to a number of sub-arguments. There were more of these than usual, and many candidates correctly identified at least one, although not many spotted three. The most popular correct answer was "Prison punishes the wrong people."

(c) As indicated in the mark scheme, this passage contained quite a few significant flaws, and several candidates spotted one or more of them. As in previous sessions, many candidates gave a running critical commentary on the ideas in the passage, instead of evaluating the reasoning; this limits their mark to 1 out of 5. Analysis, paraphrase and literary evaluation are not credited at all. Criticisms which are not credited include the absence of statistics and being one-sided (it is not the author's job to argue against his own opinions). In general, candidates should comment on what the author has said, rather than suggesting what he should have said. If teachers and candidates study mark schemes from previous sessions, they will see the kinds of answers to 3c which are expected. They can expect to find weaknesses and flaws in these passages, which are written specifically to provide opportunities for comments of that kind. Candidates who correctly identify and briefly explain two or three of these will be awarded 5 marks out of 5. A number of candidates mistakenly interpreted the passage as arguing in favour of capital punishment for some or all crimes.

(d) Although the instructions for this question make it quite clear both that the conclusion of candidates' answers must be stated and what that conclusion should be, quite a lot of candidates lost marks by presenting a different conclusion or none at all. A number of candidates gave developed personal reflections on the theme without presenting them in the form of an argument: although some of these were quite impressive, they were limited to 1 mark, because the question asked for an argument (i.e., a conclusion supported by reasons). Some candidates misunderstood the claim and challenged it on the grounds that different crimes should receive different penalties. Others mistakenly based their discussion on the ideas in the passage, instead of responding to the claim in its own terms. Some candidates assumed that the claim was asking for current penal policy to be made harsher, and some of them retorted that it was harsh enough already, which was actually not incompatible with the claim. Others thought they were being asked to discuss the imposition of capital punishment for some or all crimes.

THINKING SKILLS

Paper 9694/22
Critical Thinking

Key messages

Some candidates need to understand that expressing opinions about the issues raised or showing further knowledge of them is not the focus of the paper and cannot receive much credit if any. Some candidates also spend too much of their time re-iterating what is in the sources and this also cannot receive any credit, apart from **Question 3a** and **b** where they are required to stick closely to the text.

General comments

There were fewer questions requiring an assessment of both sides of a position or argument than perhaps is usual. Most candidates seemed to respond to the issues raised by the questions and were able to cope with the content of the sources. Able candidates were able to use the material as a vehicle for illustrating their thinking skills.

Comments on specific questions

Question 1

(a) Candidates mainly did this well and saw that the main source of reliability was the corroboration received from Sources C and E. Some candidates speculated that Mr Smith was from a rival firm but such a speculative approach cannot be credited. Equally, 'ability to see' cannot be used as a criterion if it simply means "he/she was there". Candidates also pointed out that this reliability was somewhat compromised by his vested interest in giving an exaggerated account, though very few applied this to the specific point about the tile "just missing" his son.

(b) There was a tendency for candidates simply to reiterate what was in the source and this was not enough to explain why it was useful. Most candidates who successfully moved on to usefulness pointed out that it bolstered Mr Smith's account. Other valid points made were that it suggested a general culture of poor regulation and the employee was an 'insider'. Surprisingly few made the point about the 'reverse vested interest' of the employee (i.e. he was a 'whistleblower') and confused this with *lack* of vested interest. On the other side, candidates pointed out that it might not be useful because we do not know that he was employed by the firm that built Mr Smith's villa and he might have been told to exaggerate for the purposes of creating exciting television.

(c) There was a wide range of possible answers here. Some candidates dealt with information already given in the sources and could not get any credit.

(d) The issues raised by the sources seemed to engage the interest of the candidates and a reasonable number considered plausible alternative scenarios. This enabled more candidates to access the Level 3 mark band. However, there were still a significant number of candidates who developed only one side, usually that the director was to blame. There were, as in previous sessions, examples of candidates writing less than they did for the other parts of the question in spite of this question being worth 6 marks. However, there seemed to be fewer examples of candidates getting 'carried away' on this question and writing at too great a length. Some candidates became over-speculative and went beyond anything the sources suggested but this was less in evidence than in previous sessions.

Question 2

(a) Very few candidates drew the conclusion that the evidence suggested that professional drivers did not use screen additive because, if they did, they would not be at risk of getting legionnaire's disease, no matter how much they drove and how often they used their screen-washers. Many candidates seemed to think that what was required was to re-iterate what the passage said. In this sort of question, candidates are expected to draw an inference from what they have been told that has not been drawn in the passage.

(b) This was done well with many candidates successfully explaining why these hand held washers did not replicate the conditions for the production of legionnaire bacteria in spite of their potential to do so, at first sight.

(c) A number of candidates were thrown by the statement in Source A that "the bacteria are always present in the environment" and took this to mean that bacteria would descend with the rain and wrongly concluded that this was a reliable conclusion. Other candidates dealt separately and at length with the evidence in the sources. If a question refers to evidence in two sources in this way, this should not be taken to indicate that a detailed examination of both sources is required. The intention is that candidates collate the evidence in both sources and then draw a conclusion from it. Some candidates were somewhat distracted by an extensive knowledge of rainfall, which they used at length. Candidates should only use the information they are given in the sources – it is not the aim of the syllabus to reward specialist subject knowledge which happens to pertain to the issues raised by the sources.

(d) As usual, few candidates reached Level 3, but more of them were reaching Level 2. Candidates tended to limit their answers to the drivers and not move beyond to the implications for wider incidents of legionnaire's disease. Candidates seemed to understand the issues raised by the sources.

Question 3

(a) Many candidates correctly identified the conclusion and gained 2 marks. On the other hand, this proved challenging enough for a significant minority who tended to put "gambling is essential for success".

(b) This proved more challenging, with only a minority successfully identifying 3 reasons.

(c) The crucial thing in this question is to *evaluate the reasoning* rather than to *challenge the propositions* that constitute that reasoning. Most candidates tended to do the latter typically, suggesting that gambling was very bad for you and illegal in many countries. Even where there is extended counter-argument of this type, candidates cannot move beyond Level 2 in the mark scheme. A number of candidates do not understand assumptions and refer to explicit statements in the text as 'assumptions'. Assumptions are essentially things that the author has taken for granted and might not be, on closer examination, things that can be taken for granted. Candidates need to make sure they are engaged in the right sort of exercise here. Equally, there was an increase in candidates doing genuine evaluation with some good answers exploring the equivocation in the use of the word 'gamble'. This meant there were more Level 3 answers than in past sessions.

(d) More or less split between candidates arguing for and against the proposition. Most candidates had a clear understanding of the proposition, though a minority thought it referred specifically to gambling rather than risk in general. There were some rather lengthy essay-type answers but these usually managed to have sufficient coherence to reach Level 3.

THINKING SKILLS

Paper 9694/23
Critical Thinking

Key messages

Candidates need to recognise what kind of answer is implied by each type of question and to answer accordingly. The best way of developing this skill is by studying previous question papers, mark schemes and Principal Examiner's reports.

In **Questions 1a, 1b, 1c, 2a, 2b** and **2c**, there are usually (but not invariably) two sides to the issue. In such cases, the mark is usually capped if candidates recognise only one side. They should be advised to look for two sides, but to be aware that in a few cases the correct answer may be "very useful", "not useful at all", or the equivalent.

It is almost always a good strategy to devote some time to thinking before beginning to write.

General comments

There was little or no evidence that any candidates had declined to engage with the examination in general. The lowest marks were awarded because of unclear thinking or claims being made without support.

There were very few "No Responses" and no instances of candidates apparently not having time to finish the paper.

As in previous sessions, strengths and weaknesses tended to be typical of particular Centres.

Comments on specific questions

Question 1

Unusually, this question referred to an actual historical debate, namely the fate of 'the princes in the Tower'. However, all the information required to answer the questions was provided in the sources. A few candidates were confused, and made suggestions such as that Sir John Tyrell was trying to protect Richard III (who was dead), that Richard III and Henry VII killed the princes together, or that Henry VII would not want to kill the princes because he was their father.

(a) Many candidates scored 2 marks out of 3 by noting that Tyrell's evidence was unreliable because he would be likely to say whatever his torturers wanted him to say. A significant minority claimed that the evidence was reliable, because under torture people always tell the truth; this answer was not credited. Some candidates, but not very many, achieved the third mark, by pointing out that Tyrell had excellent ability to know what had happened.

(b) Most candidates gave very little credence to the Richard III Society, on the basis that it was biased and/or lacked expertise. Relatively few gained the third mark, by pointing out that it is an academic society, which engages in serious research, or that their views were useful because they provided an alternative perspective on the history. Quite a lot of candidates unreasonably claimed that the society's views were not based on evidence, and this was not credited.

(c) There were several valid comments which could be made in response to this question, and most candidates spotted at least one of them. Some over-stated the significance, by claiming that the discovery of the skeletons "proved" that the princes had been murdered.

(d) As usual, in order to achieve Level 3, candidates needed both to evaluate the sources and to consider a possible alternative scenario. Slightly more candidates did this than in previous sessions. On this occasion, the sources themselves clearly hinted at several alternative scenarios, and most candidates discussed, or at least mentioned, one or more of them. Unsurprisingly, the most popular answers were that the princes were killed by Richard III or by Henry VII. Some candidates argued that the princes were murdered, without considering who might have been the murderer.

Question 2

This question was based on a recent genuine research project. The statistics in Source C and the information in Sources B and D were all taken from the project report. The information about physiognomy in Source A was also genuine.

(a) Most candidates achieved 2 marks by showing that the statistics in Source C did support the claim in Source D. There were several ways of achieving a third mark, and some candidates did so. A few candidates correctly and perceptively stated that the statistics strictly supported only a modified version of the claim, namely that unattractive people were more likely to *admit* to having committed crimes.

(b) Only a minority of candidates correctly identified the key point, namely that the claim was not a valid criticism of the research exercise, because allowance had been made for factors which might indicate poverty. However, there were several other valid comments which scored 1 mark, and quite a lot of candidates identified at least one of them.

(c) Not many candidates identified the key point, namely that comments about arrests and convictions were not directly relevant, because the research data were self-reported. However, there were several other valid points that could be made, and a fair number of candidates achieved at least 1 mark.

(d) Most candidates did make some reference to the source material, and a good many of them made some relevant criticisms. Very few discussed the transition from the claim of the research (that unattractive people are more likely to commit crime than attractive people) to the claim in the question (that it is possible to recognise potential criminals by their faces).

Question 3

In general, this argument made some reasonable points, but most candidates considered the conclusion to be over-stated.

(a) Most candidates, although by no means all, correctly identified the main conclusion of the argument. Although it was fairly prominent, it is unlikely that anyone would have guessed this as the right answer, rather than identifying it by analysing the argument correctly. A few candidates were awarded 1 mark for paraphrasing the conclusion or incorporating additional material.

(b) As usual, the correct answers to this question were intermediate conclusions to a number of sub-arguments. There were more of these than usual, and many candidates correctly identified at least one, although not many spotted three. The most popular correct answer was "Prison punishes the wrong people."

(c) As indicated in the mark scheme, this passage contained quite a few significant flaws, and several candidates spotted one or more of them. As in previous sessions, many candidates gave a running critical commentary on the ideas in the passage, instead of evaluating the reasoning; this limits their mark to 1 out of 5. Analysis, paraphrase and literary evaluation are not credited at all. Criticisms which are not credited include the absence of statistics and being one-sided (it is not the author's job to argue against his own opinions). In general, candidates should comment on what the author has said, rather than suggesting what he should have said. If teachers and candidates study mark schemes from previous sessions, they will see the kinds of answers to 3c which are expected. They can expect to find weaknesses and flaws in these passages, which are written specifically to provide opportunities for comments of that kind. Candidates who correctly identify and briefly explain two or three of these will be awarded 5 marks out of 5. A number of candidates mistakenly interpreted the passage as arguing in favour of capital punishment for some or all crimes.

(d) Although the instructions for this question make it quite clear both that the conclusion of candidates' answers must be stated and what that conclusion should be, quite a lot of candidates lost marks by presenting a different conclusion or none at all. A number of candidates gave developed personal reflections on the theme without presenting them in the form of an argument: although some of these were quite impressive, they were limited to 1 mark, because the question asked for an argument (i.e., a conclusion supported by reasons). Some candidates misunderstood the claim and challenged it on the grounds that different crimes should receive different penalties. Others mistakenly based their discussion on the ideas in the passage, instead of responding to the claim in its own terms. Some candidates assumed that the claim was asking for current penal policy to be made harsher, and some of them retorted that it was harsh enough already, which was actually not incompatible with the claim. Others thought they were being asked to discuss the imposition of capital punishment for some or all crimes.

THINKING SKILLS

Paper 9694/31
Problem Analysis and Solution

Key messages

This paper required candidates to engage in the full range of problem-solving skills: the questions involved careful reading of the text laying out the problems' structures, some experimental investigation of the options, and considered reflection on what best fitted the questions' requirements.

General comments

The majority of candidates made an attempt at all four questions, although many did not appreciate important aspects of the problems' structures, and hence gathered very few marks. In these cases, it was neat and careful working which distinguished those who were able to access partial credit from those who scored nothing. Although most questions could be answered by a single numerical answer, three questions required more developed solutions, and these were marred by unclear expression and layout, as well as difficulty in facing the logical challenges that such questions posed.

Comments on specific questions

Question 1

This question required candidates to have a basic understanding of probability, and could be tackled using an orderly approach to the possible quantity and timing of Offa's decree. A significant number of candidates did not manage to appreciate how the loan system worked, and were not able score. A small number of these confused themselves with the timing of the loan payment (does the half penny need to be paid before the loan is granted?). This query was answered by Ethelred's initial idea (which would have been impossible to accomplish, if the payment was required before the loan was granted), and shows how such examples provided should be used in clarifying the problems' details.

(a) Most candidates succeeded in answering both of these questions correctly, or neither. Common misconceptions included those who thought that Offa might have decreed more than one increase, and those who thought that Ethelred would take out four loans regardless of when the increase happened. Both of these starkly show the care that needs to be trained on such expressions as "…on **one** of the next four nights…".

(b) (i) There were many creative strategies suggested here, but the question was looking for something akin to that offered by Ethelred. Many candidates offered solutions in which there was a (fairly low) chance of a loss – not appreciating the categorical nature of the requirement, "…**ensuring** the she will not lose any money."

(ii) and (iv)
Both questions were based on careful enumeration of outcomes: some candidates were still offering descriptive probabilities ("pretty high"), and others attempted to use probability rules (such as the product rule). In these sorts of questions, it is almost always best to begin by listing the outcomes – and then to invoke more abstract principles if this approach is looking too laborious. Answers involving rounded decimals were tolerated; however, fractions are encouraged, when giving probabilities, since they leave less room for calculation errors and misjudgements regarding appropriate precision.

(iii) A few candidates saw that this was a very similar situation to **(a)(i)**, in which the best increase on the night of the first loan gives a maximum.

(c) This question attracted a number of solutions which exercised the rules of probability, with variable success. Partial marks were available for this. A small number of solutions were undone by candidates' weaknesses in manipulating fractions (which should not be an issue since calculators are allowed).

(d) This was an example of a question which was highly accessible, and not dependent on candidates' mastery of previous parts. A number of candidates gave up by this stage, appearing to have lost their way amidst the probability questions. In this case, careful reading of the question was all that was required, to appreciate that the 'obvious answer' was the right one.

Question 2

This question required candidates to understand the interplay between two patterns of letters (the keyboard, and the alphabet grouped in nines/sevens). The question gave no immediate instructions as to how this was to be done, and required candidates to choose adjacent letters from the keyboard, and decide whether their interchange was detectable. About three quarters of the candidates managed to accomplish this process, and gain at least one mark.

(a) (i) There was some confusion over what was required to give "three pairs of letters", with many candidates offering three triplets (such as GPY). A number of candidates appreciated that G and Y were appropriate, and managed to extend the logic to A and S, but were not able to identify the last pair.

(ii) This question was dependent on a correct solution to **(a)(i)**, and was dealt with less successfully as a result. Solutions that referred to letters rather than countries were tolerated, although candidates should be wary of using such shorthand in general. In this case, it was unambiguous and deemed a minor error. Using letters to represent countries in a different question may yield an ambiguous answer, and thus be deemed inadmissible.

(b) (i) This question was answered successfully by a small number of candidates, although it was independent of the previous answers, and did not require an understanding of the link between the two grids of letters. The three pairs of letters which were not detectable (in **(a)(i)**) required only three letters to be omitted. A small number of candidates gave the 'reversed answer' (three) – which was only credited if it was clear that these were *not* to be included. Candidates should be encouraged always to offer some unit/explanatory phrase with any numerical answer, to ensure clarity.

(ii) A small number of candidates managed to select three appropriate letters here – solutions which selected three from an incorrect answer to **(a)(i)** were credited in this case.

(c) This question required a robust appreciation of what the rows and columns of the new table represented: a number of candidates appeared to confuse rows and columns, with 'M' being a popular wrong answer. Candidates needed to return to the explanatory note above the similar table representing multiples of 9, in order to grasp the significance of the columns.

(d) The final question required a strategy, and an organised answer. The clue given in the question, stating that Q, W and E needed to be in different groups, was picked up by a few candidates. Most candidates, however, simply listed the letters haphazardly, if they managed to answer it at all. Such guesswork may often be worth a try, but was very unlikely to gain marks here.

Question 3

This question required candidates to handle averages, percentages, costs and profits with confidence. The investigative part of the question (which identifies it as a modelling question on this paper) was manageable with algebra, although this was by no means vital to achieving a solution. Many candidates made a creditable attempt at **(d)**, even though they had not mastered **(b)** and **(c)**, as is normally encouraged in such questions.

(a) (i) Most candidates appreciated that profit was to be found by subtracting the cost from the selling price. Roughly two thirds of the candidates managed to assemble the appropriate details (requiring a little care in deducing that 2000 barrels of 'light' were produced), and gained at least one mark.

(ii) This question was fairly well answered, although this was perhaps related to the small group of possible answers (0, 50, 75 and 100). Very few candidates gained 1 mark here, since those that embarked upon the appropriate working tended to get the question right.

(b) This question built substantially on the example offered in **(a)**, and required care and organisation if tackled by trial and improvement. Each trial required a couple of intricate calculations, and it was far from obvious how large the 'area' to be searched was. Very few candidates managed a successful search. About 10% of candidates managed to express the key part of the problem algebraically, using an expression such as $(6000x + 20\,000) - (8000 \quad 11)$. However, very few of these managed to formulate the appropriate inequality and solve it.

(c) This question required similar skills as those employed in **(b)**, and was met with similar levels of success.

(d) Although this question was difficult to answer without the boundaries required in **(b)** and **(c)**, many candidates attempted it, having made intelligent guesses. The most popular answers were based on 50% light and 50% heavy being produced, which gained 2 marks (out of the 5 available).

Question 4

This question required candidates to investigate the possible scores, while adhering to the multiple restrictions laid out in the rules. It required candidates to submit lists of options/scores in almost all the subsections of the question, and was more a test of candidates' ability to offer a *complete* answer than their ability to offer a possible answer. The overwhelming majority of candidates were able to answer some parts of the question correctly.

(a) The precise five correct scorelines were needed to gain full marks here; many candidates either omitted one, or included the two forbidden ones involving '7'. Some candidates struggled with the distinction between an individual score, and a scoreline. Such subtle distinctions are often drawn in these questions; candidates are advised to study all the information given carefully and look for clarifying examples.

(b) This question was well answered by most candidates. The fact that the question stated that four different numbers were possible made this easier than most of the other 'listing' questions.

(c) The need to consider two turns by each player required organisation and care here, and this question was not well answered overall. Very few candidates attempted to lay out the options systematically, or even eliminate the numbers which were impossible.

(d) This adaptation to the rules did not pose a problem to most candidates, and was well answered overall.

(e) The tracking of the scores over an entire game proved too much for most candidates, and few managed to get both possible scorelines correct. This was seemingly the result of the calculations appearing to be sufficiently easy to perform in one's head, and so a lack of working was seen; time pressure may also have been a factor in limiting what candidates recorded. Given that diligent time-consuming care is generally the most important skill required to tackle **Question 4** successfully, future candidates might consider dealing with this question first.

(f) Many candidates saw that the unlikely choice by Fay led to a forced move by Simon, and correctly described this.

(g) As with **(e)**, this required the consideration of a complete game, and was attempted by many without working, with limited success.

THINKING SKILLS

Paper 9694/32
Problem Analysis and Solution

Key messages

This paper required candidates to engage in the full range of problem-solving skills: the questions involved careful reading of the text laying out the problems' structure, some experimental investigation of the options, and considered reflection on what best fitted the questions' requirements.

General comments

The majority of candidates made an attempt at all four questions, although many did not appreciate important aspects of the problems' structures, and hence gathered very few marks. In these cases, it was neat and careful working which distinguished those who were able to access partial credit from those who scored nothing.

Comments on specific questions

Question 1

This question depended upon a careful initial assessment of what restrictions the adjacent page numbers placed on the map's layout. This was defined by the double-page map given at the beginning of the question, and was not appreciated by most candidates. It appeared that many candidates did not distinguish the necessary relations defined by the map, from the example of how the pages might be laid out, resulting in many confused solutions.

(a) (i) This question required candidates to appreciate the significance of the 14-page difference between the pages given and the pages above. Without this insight, approaches to the entire question were likely to score few marks. Some candidates did sketch out the pages (as was suggested by the example) and managed many of the ensuing questions successfully as a result.

(ii)(iii) These questions required candidates to appreciate that the pages given could have been near the left-hand or right-hand edge of the row. As with **(a)(i)**, this was helped by a sketch of how the pages were laid out. Very few candidates managed to accomplish this successfully.

(b) This question depended upon an attempt to establish how many rows could be fitted between the pages given and the beginning of the book. Many candidates attempted this question without a sketch, and struggled to establish how many rows were possible.

(c) Once again, many candidates took the example that was given (showing a stepped reduction) as defining precise new restrictions (5 rows, with 12 pages in the middle row): the answer '36' was a popular wrong answer as a result (being the number of pages in the example). The importance of this distinction (between examples which illustrate the structure, and restrictions which are precise and necessary) is vital for all such questions. As with **(a)** and **(b)**, approaches to this question were helped by a sketch of how the pages fit together.

(d) This question was difficult to approach without a sketch of how the stepped reduction of the pages affected the numbering. As mentioned earlier, some clear working of this kind was also necessary to gain any partial marks, for those who had misunderstood the significance of the example given above part **(c)**.

Question 2

This question required candidates to understand and manage the two rates of coverage given in the stem, and also to marshal the cost calculations with due care. About three quarters of the candidates managed to accomplish the second of these processes at a basic level, and gain at least one mark. The difficulty of representing the state of the arches at any time made dealing with the two paints unmanageable for many candidates; finding an appropriate method of representation is often the key to solving these sorts of questions.

(a) This question required a confident appreciation of what work was required on the arches, on any given year, when the original paint was being applied. As stated in the first paragraph, all four painters were needed full-time, without a break, in order to avoid the bridge becoming unsafe. As a result, there would be four arches requiring paint when Albert's colleagues were sacked. Surprisingly few candidates had the confidence to say that the bridge became unsafe immediately.

(b) This question was a test of candidates' organisation, in calculating the costs of the different constituent parts. Many candidates assumed that the 12 painters were being employed *instead* of the 4 original painters: candidates must be extremely vigilant for such potential misinterpretations, and then slow down in order to ensure they have assessed the logic of the question's requirements correctly.

(c) A similar set of skills were required in answering this question as for **(b)**, and many candidates did manage to calculate the cost of four painters using the new paint ($372 000). Candidates should appreciate the high likelihood of making small errors in a question like this, and set out their working to minimise the chances of this.

(d)(i) This question required imaginative experimentation with different quantities of the paints, and confidence in knowing which arches needed paint when. Very few candidates managed to establish the optimal answer. Unfortunately, other plausible answers were frequently supported by careless working, which earned the solutions no marks at all.

(ii)(iii) These two questions depended on a feasible answer to **(d)(i)**, and were successfully dealt with by only the most astute candidates.

Question 3

This question required candidates to develop and compare linear models, and was well answered by the majority of candidates. Candidates clearly felt at ease finding the linear rates of change between the drivers' data points, and many were able to formulate these algebraically (although it was not necessary for a complete solution).

(a) The vast majority of candidates were able to find the takings of each of the drivers, and identify who took the most. Some candidates were unsure whether the question required a total for all the drivers, or just the driver with the largest takings, and sensibly gave both figures. If candidates do find themselves doing this, they must ensure they explain what the figures mean (e.g. "total for Ephron = $350, total for all drivers = $1105").

(b) Elementary trial and improvement revealed a maximum here, and most candidates had no difficulty in finding it.

(c) This question required candidates to apply the process of linear extrapolation, which had just been demonstrated, to two new data points. Most candidates found this manageable, helped by the fact that no great search was needed to find a maximum.

(d) This question required candidates to explore the range of possible pairings, and efficiently dismiss unlikely ones, in order to focus on pairings which might lead to a maximum at 35 passengers. There were a number of orderly and productive solutions to this question, although these tended to treat each of the 10 possible pairings as equally worthy of investigation, while some of them could be quickly dismissed (it could quickly be seen that A&B could not maximise passengers at 35, for instance). The combination of organisation in the search, and reflective narrowing of the field of search, is what is needed in these modelling questions.

(e) A number of candidates realised that a correct answer to **(d)** was not necessary to progress to the final part. And many realised that the question boiled down to justifying which pairing would lead to 26 passengers at $27. A number of the pairings had been fully explored by this point, and quite a good proportion of solutions correctly identified Ephron as the other driver. For full marks, some attempt was needed to justify this – such as an appeal to the two figures being halfway between 45 passengers at $4, and 7 passengers at $50 – and this demonstration was not well supplied, in general. When candidates are asked to demonstrate a claim is correct, they must consider what extra information/insights will make the claim's truth clearer. Two marks were available for this here, and so a precise additional piece of information was expected.

Question 4

This question required candidates track passengers between the mismatched timetables of the ferry and the train, while abiding by the multiple constraints on price. It was designed to stretch the capacity of candidates to navigate and organise complex data, and, as with many such questions, its solution depended upon a clear laying-out of the constituent parts. The fact that an amount of working could be done on the question paper (which is permitted), should not obscure the fact that candidates can only gain partial marks if they leave a clear trace of the decisions they have made on their answer paper.

(a) This was well tackled in general – as mentioned above, candidates should be alert to the possibility of misinterpretation here (calculating the number of journeys to Peladot and back, for instance) and pick their way carefully through the requirements.

(b) This was answered fairly well – although a substantial minority either did not appreciate that a 17 year old was an adult according to the pricing system, or overlooked the fact that the cost of the car included the driver.

(c) (i) This question required candidates to tally the ferry fares, and select an appropriate train fare – which most candidates did successfully. Some candidates found the expressions 'single' and 'return' confusing, assuming that travellers needed one of each in order to get to Craugastor and back again, when in fact a 'return' ticket covered both journeys. Allowances were made in the marking for any such misunderstanding.

(ii) A correct answer to this question depended on an appreciation of which trains were available to Mrs Mander with an off-peak ticket: in this case the first train she could take was at 09:36. Some candidates decided to round the times they gave, which is not recommended in general. If train times are given to the nearest minute, then journey times can and should be given to the nearest minute as well.

(d) (i)(ii) Both of these questions required careful working backwards and forwards from given times in the question (the earliest ferry being 06:45 on a Thursday). As mentioned earlier, they were designed to stretch candidates' mental capacities to the limit, and it is recommended that times are jotted down, with reference to what they mean, in order to navigate such questions with confidence.

(e) This question required imaginative use of the ferry ticket discount system, as well as a confident understanding of the train fares. Few candidates managed to complete it successfully. This may have partly been due to time constraints. Most of those who did attempt this question managed to leave a comprehensible break-down of Jeremy's costs, and thus gained partial marks.

THINKING SKILLS

Paper 9694/33
Problem Analysis and Solution

Key messages

This paper required candidates to engage in the full range of problem-solving skills: the questions involved careful reading of the text laying out the problems' structures, some experimental investigation of the options, and considered reflection on what best fitted the questions' requirements.

General comments

The majority of candidates made an attempt at all four questions, although many did not appreciate important aspects of the problems' structures, and hence gathered very few marks. In these cases, it was neat and careful working which distinguished those who were able to access partial credit from those who scored nothing. Although most questions could be answered by a single numerical answer, three questions required more developed solutions, and these were marred by unclear expression and layout, as well as difficulty in facing the logical challenges that such questions posed.

Comments on specific questions

Question 1

This question required candidates to have a basic understanding of probability, and could be tackled using an orderly approach to the possible quantity and timing of Offa's decree. A significant number of candidates did not manage to appreciate how the loan system worked, and were not able score. A small number of these confused themselves with the timing of the loan payment (does the half penny need to be paid before the loan is granted?). This query was answered by Ethelred's initial idea (which would have been impossible to accomplish, if the payment was required before the loan was granted), and shows how such examples provided should be used in clarifying the problems' details.

(a) Most candidates succeeded in answering both of these questions correctly, or neither. Common misconceptions included those who thought that Offa might have decreed more than one increase, and those who thought that Ethelred would take out four loans regardless of when the increase happened. Both of these starkly show the care that needs to be trained on such expressions as "…on **one** of the next four nights…".

(b) (i) There were many creative strategies suggested here, but the question was looking for something akin to that offered by Ethelred. Many candidates offered solutions in which there was a (fairly low) chance of a loss – not appreciating the categorical nature of the requirement, "…**ensuring** the she will not lose any money."

(ii) and (iv)
Both questions were based on careful enumeration of outcomes: some candidates were still offering descriptive probabilities ("pretty high"), and others attempted to use probability rules (such as the product rule). In these sorts of questions, it is almost always best to begin by listing the outcomes – and then to invoke more abstract principles if this approach is looking too laborious. Answers involving rounded decimals were tolerated; however, fractions are encouraged, when giving probabilities, since they leave less room for calculation errors and misjudgements regarding appropriate precision.

(iii) A few candidates saw that this was a very similar situation to **(a)(i)**, in which the best increase on the night of the first loan gives a maximum.

(c) This question attracted a number of solutions which exercised the rules of probability, with variable success. Partial marks were available for this. A small number of solutions were undone by candidates' weaknesses in manipulating fractions (which should not be an issue since calculators are allowed).

(d) This was an example of a question which was highly accessible, and not dependent on candidates' mastery of previous parts. A number of candidates gave up by this stage, appearing to have lost their way amidst the probability questions. In this case, careful reading of the question was all that was required, to appreciate that the 'obvious answer' was the right one.

Question 2

This question required candidates to understand the interplay between two patterns of letters (the keyboard, and the alphabet grouped in nines/sevens). The question gave no immediate instructions as to how this was to be done, and required candidates to choose adjacent letters from the keyboard, and decide whether their interchange was detectable. About three quarters of the candidates managed to accomplish this process, and gain at least one mark.

(a) (i) There was some confusion over what was required to give "three pairs of letters", with many candidates offering three triplets (such as GPY). A number of candidates appreciated that G and Y were appropriate, and managed to extend the logic to A and S, but were not able to identify the last pair.

(ii) This question was dependent on a correct solution to **(a)(i)**, and was dealt with less successfully as a result. Solutions that referred to letters rather than countries were tolerated, although candidates should be wary of using such shorthand in general. In this case, it was unambiguous and deemed a minor error. Using letters to represent countries in a different question may yield an ambiguous answer, and thus be deemed inadmissible.

(b) (i) This question was answered successfully by a small number of candidates, although it was independent of the previous answers, and did not require an understanding of the link between the two grids of letters. The three pairs of letters which were not detectable (in **(a)(i)**) required only three letters to be omitted. A small number of candidates gave the 'reversed answer' (three) – which was only credited if it was clear that these were *not* to be included. Candidates should be encouraged always to offer some unit/explanatory phrase with any numerical answer, to ensure clarity.

(ii) A small number of candidates managed to select three appropriate letters here – solutions which selected three from an incorrect answer to **(a)(i)** were credited in this case.

(c) This question required a robust appreciation of what the rows and columns of the new table represented: a number of candidates appeared to confuse rows and columns, with 'M' being a popular wrong answer. Candidates needed to return to the explanatory note above the similar table representing multiples of 9, in order to grasp the significance of the columns.

(d) The final question required a strategy, and an organised answer. The clue given in the question, stating that Q, W and E needed to be in different groups, was picked up by a few candidates. Most candidates, however, simply listed the letters haphazardly, if they managed to answer it at all. Such guesswork may often be worth a try, but was very unlikely to gain marks here.

Question 3

This question required candidates to handle averages, percentages, costs and profits with confidence. The investigative part of the question (which identifies it as a modelling question on this paper) was manageable with algebra, although this was by no means vital to achieving a solution. Many candidates made a creditable attempt at **(d)**, even though they had not mastered **(b)** and **(c)**, as is normally encouraged in such questions.

(a) (i) Most candidates appreciated that profit was to be found by subtracting the cost from the selling price. Roughly two thirds of the candidates managed to assemble the appropriate details (requiring a little care in deducing that 2000 barrels of 'light' were produced), and gained at least one mark.

(ii) This question was fairly well answered, although this was perhaps related to the small group of possible answers (0, 50, 75 and 100). Very few candidates gained 1 mark here, since those that embarked upon the appropriate working tended to get the question right.

(b) This question built substantially on the example offered in **(a)**, and required care and organisation if tackled by trial and improvement. Each trial required a couple of intricate calculations, and it was far from obvious how large the 'area' to be searched was. Very few candidates managed a successful search. About 10% of candidates managed to express the key part of the problem algebraically, using an expression such as $(6000x + 20\,000) - (8000 \times 11)$. However, very few of these managed to formulate the appropriate inequality and solve it.

(c) This question required similar skills as those employed in **(b)**, and was met with similar levels of success.

(d) Although this question was difficult to answer without the boundaries required in **(b)** and **(c)**, many candidates attempted it, having made intelligent guesses. The most popular answers were based on 50% light and 50% heavy being produced, which gained 2 marks (out of the 5 available).

Question 4

This question required candidates to investigate the possible scores, while adhering to the multiple restrictions laid out in the rules. It required candidates to submit lists of options/scores in almost all the subsections of the question, and was more a test of candidates' ability to offer a *complete* answer than their ability to offer a possible answer. The overwhelming majority of candidates were able to answer some parts of the question correctly.

(a) The precise five correct scorelines were needed to gain full marks here; many candidates either omitted one, or included the two forbidden ones involving '7'. Some candidates struggled with the distinction between an individual score, and a scoreline. Such subtle distinctions are often drawn in these questions; candidates are advised to study all the information given carefully and look for clarifying examples.

(b) This question was well answered by most candidates. The fact that the question stated that four different numbers were possible made this easier than most of the other 'listing' questions.

(c) The need to consider two turns by each player required organisation and care here, and this question was not well answered overall. Very few candidates attempted to lay out the options systematically, or even eliminate the numbers which were impossible.

(d) This adaptation to the rules did not pose a problem to most candidates, and was well answered overall.

(e) The tracking of the scores over an entire game proved too much for most candidates, and few managed to get both possible scorelines correct. This was seemingly the result of the calculations appearing to be sufficiently easy to perform in one's head, and so a lack of working was seen; time pressure may also have been a factor in limiting what candidates recorded. Given that diligent time-consuming care is generally the most important skill required to tackle **Question 4** successfully, future candidates might consider dealing with this question first.

(f) Many candidates saw that the unlikely choice by Fay led to a forced move by Simon, and correctly described this.

(g) As with **(e)**, this required the consideration of a complete game, and was attempted by many without working, with limited success.

THINKING SKILLS

Paper 9694/41
Applied Reasoning

Key message

Candidates need to take time to read each question carefully, and ensure they have fully understood the demands of the question before they attempt to respond. Several obviously able candidates did not do full justice to themselves, particularly on **Questions 1(b)** and **4**, through not responding to the specifics of these questions, which they overlooked or misconstrued.

General comments

It was good to see many candidates appropriately responding to **Question 4**, applying thinking skills to current affairs. The overall performance though was impacted by the generally poorer performance in **Questions 2** and **3**, where a great many candidates did not demonstrate knowledge of the structure of arguments, or competence with the language of reasoning, using the skills CT4 and CT6, as prescribed by the syllabus.

Comments on specific questions

Question 1

Candidates needed to understand that while **(a)** requires a distinctively different answer to **(b)**, that answer should not contradict anything they said in **(a)**. Most candidates answered **1(a)** better than **1(b)**. Candidates were only required to respond very briefly, but with precision and clarity. As in previous sessions, several wrote unnecessarily discursive and long answers, up to a page, a disproportionate length for a 1 mark per criticism point in **(a)** and 2 mark brief explanation in **(b)**.

(a) In this question candidates were required to identify errors or flaws in the statistics as they were presented in the advertisement and the bar chart. Almost all candidates scored at least 1 mark and the majority got 2. It was good that a large number not only noticed the truncated scale on the *x*-axis, but were able to express the idea that the scale was deliberately misleading. A minority of candidates successfully questioned the phrase “in the last 10 years”, and there were some creditworthy references to biased presentation. Very few, though, criticised the selection of 10 years as being an inappropriate measure of truck longevity.

It is worth noting that any response along the lines of “there is no evidence for / back-up data for...” is unlikely ever to be credited.

(b) The overall performance on this question was poor. Candidates needed to have focused on the specifics of the claim only and decide if they validly described Chase. A number of candidates confused the criticism of the reliability of the data undertaken in part **(a)** with the inference about Chase as “most reliable” in **(b)**, responding that since the reliability of the data was questionable, Chase could not be the “most reliable” truck on the road. The question however required them to focus on the relevance or sufficiency of the data presented to making the stated claim. Only a minority of candidates recognised the mismatch between the data presented and the terms ‘reliable’ or ‘long-lasting’. Even fewer were able to explain the mismatch sufficiently well to be awarded both marks. A similarly small percentage explored the ‘solid as rock’ simile.

A good many candidates repeated their answers for **(a)** in **(b)**, for which they could not gain any extra credit; or were not credited because they responded in **(a)** with an answer that should have been a response to **(b)**. Some others offered responses inconsistent with the criticisms they had made in **(a)**. For example, several candidates simply took it from the visual presentation of bar chart that the given claim could be inferred, whereas in **(a)** they had identified accurately the weaknesses in the presentation. Moreover, it demonstrated a worryingly weak grasp of skill CT14 when some candidates were happy to accept that a truck brand is reliable simply because a graph showed that 1% more of its trucks are still on the road after 10 years, when compared with another brand, as part of an advertisement for a company. The syllabus requires candidates to adopt a sceptical approach to statistical inferences and query their validity.

Question 2

Candidates, on the whole, found this more challenging than in the previous session. Frequently only 1 mark was accessed, because the candidate did hardly any analysis of the structure of the argument, or haphazardly identified at least one correct feature of structure. The question requires candidates to work through the passage systematically, identifying and labelling important component parts of the argument correctly, such as the main conclusion, main reasons and intermediate conclusions, and the counter-argument. Candidates should keep as precisely with the wording in the text as possible, and avoid additional material from the text which goes beyond the core reason or conclusion identified. For example, *"These statistical snapshots are not helpful"* is a correctly-identified intermediate conclusion; many candidates went beyond that to add *"often the proffered ranking is a spurious correlation"*, which is adding more reasons and not precisely identifying what comprised the intermediate conclusion only. Candidates should also write out the element they have identified in full and not use ellipsis.

Although this mistake has been identified in several previous reports, there are still a fair number of candidates who are answering the question incorrectly by commenting on the literary features and stylistic devices of the passage. For example, *"Quotes and examples are explained in the ensuing paragraphs by giving good descriptions. Each paragraph has a point and reason that strengthens the main conclusion"*. There is no creditworthy element in such a response, as this is not doing critical thinking, but doing literary criticism. Analysing the structure of an argument is a technical exercise, requiring the candidate to uncover the skeletal framework of the reasoning which holds the argument together.

Some other candidates went on to give annotations and explanations of why they had chosen a main conclusion (MC), or an intermediate conclusion (IC) etc. For example, *"This is Hazlett's main conclusion because it conveys to the reader that...."* or *"This IC provides understanding that...."*. Such observations earn no credit.

Question 3

As has been emphasised in previous reports, this key question precisely tests crucial skills that lie at the heart of critical thinking activity – CT4 and CT6. Candidates are required to evaluate a given argument in order to come up with a judgment about the strength of the argument, and how well the main conclusion is supported. They have to do this by looking for flaws and unstated assumptions which can weaken the main conclusion, and explain each weakness they identify with precision and clarity. A small proportion of candidates identified and clearly explained the major weaknesses in the author's argument. But on the whole the question was poorly attempted, with many scoring under 3 marks. A great many candidates were unfamiliar with the critical thinking terminology such as 'unstated assumptions' (CT4), and merely made re-statements of claims in the passage, calling them 'assumptions'. Some candidates were able to pick up an occasional error in the reasoning accurately, but lacked the language of reasoning (which is acquired through practice of CT1–12) to precisely develop the point and earn two marks.

A common trend was that many candidates started this question by marking out strengths, and by doing so ended up with an appraisal of the argument as a literary artefact, and not with a critical evaluation of the reasoning. Candidates should begin by targeting obvious weaknesses that are in the argument. If they do find any strengths in these arguments, these should only be things which give very strong support to the main conclusion, such as a highly appropriate and compelling analogy, for example. A fair number of candidates offered opinions about what they thought about the argument and whether they agreed with it or not, rather than how it was flawed at points. For example, *"the outlook that we should be believing our children are lagging behind can be very dangerous since it builds over confidence"* scores nothing, as it is mere opinion, not evaluation of a weakness. Others did not score credit because they were merely challenging claims or making counter-assertions, such as *"it is wrong to say that money put into broadband would not improve the country"* or *"where is the evidence that....?"*

Question 4

The majority of candidates made a reasonable attempt of this question. Several candidates offered well-reasoned cases that ascended the scale of good applied reasoning, integrating critical inferences from the stimulus sources with examples from current affairs they were aware of. The best answers demonstrated good evaluation of sources, implicit and explicit critical comments, and juxtaposition of sources, comparing and contrasting them to forge sound critical reasoning. Responses which had good evaluation of sources, but then did not go on to produce further arguments, using their own ideas and examples, were limited to lower Band 3.

It was good that most candidates appeared aware that credit is given for reference to documents; but in referring to all or most of the documents, many candidates were not selective about which points to use in order to best support their arguments, and weaker responses only summarised the documents without being critical, serving no point. Stronger candidates selected and re-organised material from across the sources judiciously, following only the order of their own reasoning, not the order in which the source documents were presented. Such responses contained synthesis of viewpoints and outlined complexities and counter-arguments, so obviously moving up the Band 3 level, and possibly into Band 4.

A significant number of candidates wrote arguments about the wrong or incomplete conclusion – either that the US economy should be on top or that the US should innovate. Candidates who focused precisely on the question of *why the US should lead in innovation (not just innovate) if it is to keep its place as the world's biggest economy (assuming it is the world's biggest economy),* were usually able to construct relevant arguments with a complete main conclusion. A well-reasoned case had to have a conclusion that clearly embraced the full extent of the given proposition in order to access higher band marks. Even if candidates did not articulate their conclusion concisely, provided they had embraced all the elements in the proposition by spreading these elements over a compound of the main and intermediate conclusion, they were deemed to have provided a complete main conclusion.

THINKING SKILLS

Paper 9694/42
Applied Reasoning

Key message

Candidates need to take time to read each question carefully, and ensure they have fully understood the demands of each question before they attempt to respond.

General comments

Overall **Questions 2** and **4** were appropriately answered, with many responses showing awareness of the syllabus prescriptions and ability to apply critical thinking in a holistic manner. There is much room for improvement in **Question 1**, where more than half of the candidates did not appear to know precisely what is required in evaluating the validity of statistics presented, as different to evaluating the credibility of evidence. A majority of candidates needed to have explained flaws and weaknesses as required in **Question 3** with precision and clarity, rather than simply challenging Baik's claims.

Several obviously able candidates also did not do full justice to **Question 4**, through not responding to the specifics of this question, which they overlooked, thereby providing digressive or half-way responses.

Comments on specific questions

Question 1

As a general point, candidates who have more comprehensive knowledge of the topics used as source material for **Question 1** (in this case biology or medical science) should concentrate only on the problems with the statistics, and not the wider context of the subject.

The question requires candidates to identify and explain weaknesses, errors or flaws in the statistics presented and any inferences drawn with precision and clarity. Most candidates were able to get at least two marks – the most common criticism of the statistics being that Americans and Japanese had differing lifestyles, while fewer noted the more subtle weaknesses in the statistics such as the absence of mention of the number of children affected by Type 2 diabetes, which might have been insignificantly small. Many responses were unnecessarily long and detailed, with annotations and explanations that were not required. Several candidates went to unnecessary lengths and provided alternative explanations. Responses such as *"Americans eat a lot of unhealthy junk food which causes them to be obese and can cause heart-related diseases"* are not criticisms of statistics, but annotations of the text with no creditworthy element. The response that *"milk has been a staple food for children from time immemorial and is good for the bones; so it is hard to believe how it can be the cause of breast cancer now"* is not a criticism of the statistics, but challenging the credibility of the evidence, which is not the focus of this question.

There were several candidates who did not appear to understand that the question was to do with interpretation of statistics, and simply made a summary of the data.

Question 2

It was good to see a majority of candidates gaining at least 3 if not 4 marks. They worked systematically through the structure, labelling the intermediate conclusions, main reasons and main conclusion. A majority of candidates did not identify the counter-argument. Candidates should understand that the counter-argument is an integral part of an argument's structure, and not a minor or redundant feature, and should be competent in discerning it within any given argument.

Question 3

This key question tests the crucial skills of CT4 and CT6 which lie at the heart of the critical thinking activity. The overall performance for this was poorer than for **Question 2**. The best responses made sustained evaluations looking for assumptions, flaws and, lastly only, strengths, if any. A majority of candidates simply offered counter-assertions or challenged claims, rather than critically evaluating the reasoning.

Candidates need to look for errors and weaknesses of reasoning which may weaken the main conclusion of the argument. Candidates who go about looking for strengths in the argument before they have evaluated the weaknesses tended to deviate from the assignment and provide literary critical appraisals, commending the author, which could not be credited.

A number of candidates who did show familiarity with CT6 unfortunately only named flaws they had spotted. *"This is an appeal to authority"* scores no marks, because "appeal to authority" remains a mere label if the weakness is not explained.

Others identified the flaw precisely, but did not develop this into an explanation. For example, "*Appeal to authority: It cannot be concluded that the concept of democracy is unchallengeable because two leaders of two of the biggest democracies said this"* scores 1 mark for identification of flaw or limited evaluation.

However, *"It cannot be concluded that the concept of democracy is unchallengeable because two leaders of two of the biggest democracies said so. This is too strong an appeal to authority. Just because they are great leaders does not mean they cannot be mistaken."* is a developed explanation of this flaw which would earn 2 marks.

Several candidates offered strong counter-arguments, based on their feelings and opinions, in place of proper evaluation. Such responses strayed beyond the requirements of the question, or beyond the scope the information given in the stimulus passage, and so gained no marks. For example, *"When Lincoln, Churchill and Gandhi gave their definition of democracy, they did not consider the dirty politics and political games that are played in modern-day democratic elections."*

Question 4

A majority of candidates attempted this question appropriately, demonstrating ability to apply critical thinking skills to real-life situations, as relevant to the given proposition. The best responses evaluated the stimulus sources, drawing out critical inferences. These insights and intermediate conclusions were then integrated with their own ideas and relevant examples from current affairs in their own country to build a coherent case, with a full conclusion for or against the proposition.

It needs to be pointed out that what is required is for candidates not to merely comment on each source in order from documents 1–4, but to evaluate them in the context of answering the question, comparing and contrasting the corroborating or conflicting claims, constructing critical reasoning and pinpointing complexities that are evident. Candidates need to select only relevant material from across the sources, with a view to supporting or challenging reasons they offer, in the process of building a case for or against the proposition. Merely summarising each document, or agreeing with each or re-phrasing the claims within each document would limit the response to lower Band 2 marks. Some candidates gave a smattering of some independent reasoning and/or implicit critical comments, and a clear conclusion of where they stood on the debate, which carried them into upper Band 2. Candidates who made some critical evaluation of the source documents and, using precise critical comments, started building up a reasoned case, accessed marks in lower Band 3.

The best responses had taken time to read the proposition carefully, so that in building up a reasoned case they offered a main conclusion that encompassed the key elements of the proposition. Several candidates discussed at length the merits and disadvantages of democracy as compared with communism, but did not draw a full conclusion in response to the proposition. The issue is not *whether democracy was the best system and should be defended* but whether democracy should be defended <u>*at all costs*</u>. There were good critically-considered responses that did not offer a complete main conclusion – for example, *"Although democracy is precious, as shown in document 1, the time and effort required for the building of said consensus, as pointed out in document 3, renders it unproductive* (critical reasoning)*. Therefore in a changing global scenario democracy has no place to be".* This is, at best, an intermediate conclusion. Concluded as it is, it has digressed from the argument and seems to answer to a different argument – *should we do away with democracy?* Such a response cannot rise above lower Band 3. Better answers with

complete conclusion, but without a consideration of counter–positions, would have ascended up to higher Band 3 but not Band 4.

An example of a conclusion that responds fully to the proposition and has also anticipated counter-positions would be, *"Therefore we can conclude that democracy has many commendable benefits,* (counter-position) *but it does not need to be defended at all costs, since it is no sacred cow but only a man-made system to maintain stability and prevent chaos amongst the population."* Such a conclusion, provided it was based on a well-reasoned case, has responded critically to the stimulus sources, and satisfied the criteria in upper Band 3, should allow access to Band 4 marks.

THINKING SKILLS

Paper 9694/43
Applied Reasoning

Key message

Candidates need to take time to read each question carefully, and ensure they have fully understood the demands of the question before they attempt to respond. Several obviously able candidates did not do full justice to themselves, particularly on **Questions 1(b)** and **4**, through not responding to the specifics of these questions, which they overlooked or misconstrued.

General comments

It was good to see many candidates appropriately responding to **Question 4**, applying thinking skills to current affairs. The overall performance though was impacted by the generally poorer performance in **Questions 2** and **3**, where a great many candidates did not demonstrate knowledge of the structure of arguments, or competence with the language of reasoning, using the skills CT4 and CT6, as prescribed by the syllabus.

Comments on specific questions

Question 1

Candidates needed to understand that while **(a)** requires a distinctively different answer to **(b)**, that answer should not contradict anything they said in **(a)**. Most candidates answered **1(a)** better than **1(b)**. Candidates were only required to respond very briefly, but with precision and clarity. As in previous sessions, several wrote unnecessarily discursive and long answers, up to a page, a disproportionate length for a 1 mark per criticism point in **(a)** and 2 mark brief explanation in **(b)**.

(a) In this question candidates were required to identify errors or flaws in the statistics as they were presented in the advertisement and the bar chart. Almost all candidates scored at least 1 mark and the majority got 2. It was good that a large number not only noticed the truncated scale on the *x*-axis, but were able to express the idea that the scale was deliberately misleading. A minority of candidates successfully questioned the phrase “in the last 10 years”, and there were some creditworthy references to biased presentation. Very few, though, criticised the selection of 10 years as being an inappropriate measure of truck longevity.

It is worth noting that any response along the lines of “there is no evidence for / back-up data for...” is unlikely ever to be credited.

(b) The overall performance on this question was poor. Candidates needed to have focused on the specifics of the claim only and decide if they validly described Chase. A number of candidates confused the criticism of the reliability of the data undertaken in part **(a)** with the inference about Chase as “most reliable” in **(b)**, responding that since the reliability of the data was questionable, Chase could not be the “most reliable” truck on the road. The question however required them to focus on the relevance or sufficiency of the data presented to making the stated claim. Only a minority of candidates recognised the mismatch between the data presented and the terms ‘reliable’ or ‘long-lasting’. Even fewer were able to explain the mismatch sufficiently well to be awarded both marks. A similarly small percentage explored the ‘solid as rock’ simile.

A good many candidates repeated their answers for **(a)** in **(b)**, for which they could not gain any extra credit; or were not credited because they responded in **(a)** with an answer that should have been a response to **(b)**. Some others offered responses inconsistent with the criticisms they had made in **(a)**. For example, several candidates simply took it from the visual presentation of bar chart that the given claim could be inferred, whereas in **(a)** they had identified accurately the weaknesses in the presentation. Moreover, it demonstrated a worryingly weak grasp of skill CT14 when some candidates were happy to accept that a truck brand is reliable simply because a graph showed that 1% more of its trucks are still on the road after 10 years, when compared with another brand, as part of an advertisement for a company. The syllabus requires candidates to adopt a sceptical approach to statistical inferences and query their validity.

Question 2

Candidates, on the whole, found this more challenging than in the previous session. Frequently only 1 mark was accessed, because the candidate did hardly any analysis of the structure of the argument, or haphazardly identified at least one correct feature of structure. The question requires candidates to work through the passage systematically, identifying and labelling important component parts of the argument correctly, such as the main conclusion, main reasons and intermediate conclusions, and the counter-argument. Candidates should keep as precisely with the wording in the text as possible, and avoid additional material from the text which goes beyond the core reason or conclusion identified. For example, *"These statistical snapshots are not helpful"* is a correctly-identified intermediate conclusion; many candidates went beyond that to add *"often the proffered ranking is a spurious correlation"*, which is adding more reasons and not precisely identifying what comprised the intermediate conclusion only. Candidates should also write out the element they have identified in full and not use ellipsis.

Although this mistake has been identified in several previous reports, there are still a fair number of candidates who are answering the question incorrectly by commenting on the literary features and stylistic devices of the passage. For example, *"Quotes and examples are explained in the ensuing paragraphs by giving good descriptions. Each paragraph has a point and reason that strengthens the main conclusion"*. There is no creditworthy element in such a response, as this is not doing critical thinking, but doing literary criticism. Analysing the structure of an argument is a technical exercise, requiring the candidate to uncover the skeletal framework of the reasoning which holds the argument together.

Some other candidates went on to give annotations and explanations of why they had chosen a main conclusion (MC), or an intermediate conclusion (IC) etc. For example, *"This is Hazlett's main conclusion because it conveys to the reader that...."* or *"This IC provides understanding that...."*. Such observations earn no credit.

Question 3

As has been emphasised in previous reports, this key question precisely tests crucial skills that lie at the heart of critical thinking activity – CT4 and CT6. Candidates are required to evaluate a given argument in order to come up with a judgment about the strength of the argument, and how well the main conclusion is supported. They have to do this by looking for flaws and unstated assumptions which can weaken the main conclusion, and explain each weakness they identify with precision and clarity. A small proportion of candidates identified and clearly explained the major weaknesses in the author's argument. But on the whole the question was poorly attempted, with many scoring under 3 marks. A great many candidates were unfamiliar with the critical thinking terminology such as 'unstated assumptions' (CT4), and merely made re-statements of claims in the passage, calling them 'assumptions'. Some candidates were able to pick up an occasional error in the reasoning accurately, but lacked the language of reasoning (which is acquired through practice of CT1–12) to precisely develop the point and earn two marks.

A common trend was that many candidates started this question by marking out strengths, and by doing so ended up with an appraisal of the argument as a literary artefact, and not with a critical evaluation of the reasoning. Candidates should begin by targeting obvious weaknesses that are in the argument. If they do find any strengths in these arguments, these should only be things which give very strong support to the main conclusion, such as a highly appropriate and compelling analogy, for example. A fair number of candidates offered opinions about what they thought about the argument and whether they agreed with it or not, rather than how it was flawed at points. For example, *"the outlook that we should be believing our children are lagging behind can be very dangerous since it builds over confidence"* scores nothing, as it is mere opinion, not evaluation of a weakness. Others did not score credit because they were merely challenging claims or making counter-assertions, such as *"it is wrong to say that money put into broadband would not improve the country"* or *"where is the evidence that....?"*

Question 4

The majority of candidates made a reasonable attempt of this question. Several candidates offered well-reasoned cases that ascended the scale of good applied reasoning, integrating critical inferences from the stimulus sources with examples from current affairs they were aware of. The best answers demonstrated good evaluation of sources, implicit and explicit critical comments, and juxtaposition of sources, comparing and contrasting them to forge sound critical reasoning. Responses which had good evaluation of sources, but then did not go on to produce further arguments, using their own ideas and examples, were limited to lower Band 3.

It was good that most candidates appeared aware that credit is given for reference to documents; but in referring to all or most of the documents, many candidates were not selective about which points to use in order to best support their arguments, and weaker responses only summarised the documents without being critical, serving no point. Stronger candidates selected and re-organised material from across the sources judiciously, following only the order of their own reasoning, not the order in which the source documents were presented. Such responses contained synthesis of viewpoints and outlined complexities and counter-arguments, so obviously moving up the Band 3 level, and possibly into Band 4.

A significant number of candidates wrote arguments about the wrong or incomplete conclusion – either that the US economy should be on top or that the US should innovate. Candidates who focused precisely on the question of *why the US should lead in innovation (not just innovate) if it is to keep its place as the world's biggest economy (assuming it is the world's biggest economy),* were usually able to construct relevant arguments with a complete main conclusion. A well-reasoned case had to have a conclusion that clearly embraced the full extent of the given proposition in order to access higher band marks. Even if candidates did not articulate their conclusion concisely, provided they had embraced all the elements in the proposition by spreading these elements over a compound of the main and intermediate conclusion, they were deemed to have provided a complete main conclusion.