

THINKING SKILLS

Paper 9694/21
Critical Thinking

Key messages

Some candidates appeared not to know what they were expected to do in response to certain questions. Candidates and their teachers should study previous question papers, mark schemes and Examiner reports, so that they understand what kinds of answer are expected.

When answering **Questions 1d** and **2d**, candidates should make use of evaluative insights about the documents which they have already made in answer to earlier questions.

When questions ask for 'alternative' explanations for evidence, candidates need to realise that paraphrasing the explanation given in the source does not constitute an 'alternative' and will therefore not be credited.

General comments

Some candidates did not offer answers to all the questions. Some of them probably ran out of time, in which case they would have been wise to produce at least a short answer to **Question 3d**, omitting **Question 3c** if necessary in order to do so. Candidates who omitted several questions probably did so because they did not know how to approach them.

Comments on specific questions

Question 1

Virtually all candidates understood the issue, although some had difficulty with one or more of the sources.

- (a) Many candidates gave valid reasons why the report from the investigators from the World Bank (or the World Bank itself) was reliable, but a significant proportion suggested unconvincing reasons as to why it might be unreliable, while others summarised or discussed the content of the report without mentioning its reliability. Since **Question 1** on this exam always includes at least one question about reliability, candidates need to know what such questions mean and how to answer them. Some candidates unrealistically criticised the investigation for taking place six years after the granting of the loan.
- (b) The two most popular answers were also the most significant, namely that the Prime Minister might be trying to prevent his own corruption from being revealed and that he might be intending to divert the generous grant from the UN to his own coffers. As indicated on the mark scheme, other answers were also credited. Some answers which were on the right lines were judged to be incomplete and awarded only 1 mark. Several candidates suggested that the Prime Minister might sincerely desire to eradicate corruption from the country, but this was not an 'alternative' motive to the one stated in Source C, and was therefore not credited.
- (c) Some candidates seemed to find it difficult to understand why the authors of Source E regarded an increase in corruption convictions as bad news, but most succeeded in explaining that the increase could be interpreted as a measure of the success of the police, the judicial system or the work of the Anti-Corruption Commission.
- (d) Most candidates judged that corruption was not being reduced, but some argued successfully for the opposite judgement. Some candidates limited themselves to identifying relevant material from the sources and using it to support their conclusion, but a fair number accessed the higher marks



by using appropriate evaluation of sources and inferential reasoning; however, some inferential reasoning was not credited because it was stated with greater certainty than was justified. Not many identified or discussed an alternative conclusion.

Question 2

Virtually all candidates seemed comfortable with the subject matter of this question and understood the point of all the sources.

- (a) A few candidates succeeded in suggesting alternative explanations for the correlation between time spent playing video games and success in various psychological tests, but the most popular approach was to explain why playing the games would develop the relevant skills, which was not an 'alternative' explanation of the correlation and was therefore not credited. Some answers of this kind relied on a narrow interpretation of 'develops', suggesting for example that the games might improve the skills but not develop them, but this was considered to be a distinction without a difference. Other candidates apparently relied on an irrelevant distinction between manual and intellectual skills. Another unsatisfactory approach was to suggest that some people might already have the requisite skills to perform well on the tests, without mentioning why they might play more video games than other people.
- (b) Most candidates managed to say something worthy of credit in response to this question, but not many achieved the full 4 marks. The most popular answer was a combination of the first two answers on the mark scheme, which was awarded 2 marks. Many answers were along the right lines without being precise or complete and achieved 1 mark out of 2. Answers criticising the evidence for lacking statistical support were not credited.
- (c) A fair proportion of candidates correctly pointed out that the evidence referred to success on the simulator, not actual operations, and also that it related to only one kind of operation. Fewer candidates offered the other two answers. Some stated that the results applied only to medical candidates, which was significantly different from the correct version of this point (that it applied only to inexperienced surgeons) and was therefore not credited. Speculative answers, such as that the sample size may have been too small for a valid conclusion to be drawn, were not credited.
- (d) As usual, many candidates were awarded 3 marks out of 6 for summarising the sources on both sides of the debate and concluding either in favour of the claim or against it. Some candidates, however, achieved higher marks, by including evaluation of sources and inferential reasoning. Quite a lot of candidates were awarded 2 marks for 'nuanced' conclusions, such as that the playing of video games should be encouraged within reason, but not so much that it might become an addiction, or that only non-violent games should be encouraged. Many candidates who had correctly identified weaknesses in the sources when answering **Questions 2b** and **2c** accepted the claims at face value in **2d**, thereby missing an opportunity to gain marks for evaluating sources. Very few candidates pointed out that the findings in Source A are ambivalent about the benefits of playing video games.

Question 3

- (a) The most popular answer was the correct one, but many incorrect answers were also given. A very few candidates were awarded 1 mark, usually because they paraphrased the correct answer instead of quoting it *verbatim*.
- (b) Many candidates correctly identified two or three correct answers to this question. The first sentence of paragraph 5 was a popular wrong answer.
- (c) A fair number of candidates correctly identified the causal flaw in paragraph 1 and the appeal to emotion and *non sequitur* in paragraph 2. The weakness in reasoning in paragraph 3 was expressed in various ways. Some candidates criticised the generalisations in the last sentence of paragraph 1 and the third sentence of paragraph 2, but these criticisms were not credited, because these generalisations were not considered to be flawed.

As on previous occasions, a significant proportion of candidates misunderstood the nature of this task, explaining why they agreed or disagreed with all or some of the opinions and claims in the passage instead of evaluating the reasoning. Evaluative points which were not credited included criticism of the argument for being one-sided, for expressing the opinion of the implied author, or

for lacking statistical support. The question made it clear that the only answers which would be credited were “flaws, unstated assumptions and other weaknesses”, but some candidates wasted time and effort identifying strengths. Literary criticism was also not credited. Some candidates understood the significance of the expression “unstated assumption” and correctly identified one or more such assumptions which significantly weakened the reasoning (particularly the one in paragraph 4), but – as on previous occasions – some misunderstood it to mean “unsupported statement” and contradictorily claimed that quotations from the passage were unstated.

- (d) Potentially, answers to this question could have considered the philanthropic provision of good quality housing, sports facilities and social amenities for workers, but very few candidates focused their answers more broadly than health and safety. As stated in the question, no credit was given for ideas derived from the passage, and this caused a few candidates to be awarded 0 marks. A fair number of candidates produced coherent strands of reasoning and made good use of intermediate conclusions or other argument elements, particularly examples. At the other end of the range, some candidates achieved only 1 mark, because their answer did not constitute an argument, or 0 marks because their comments were either incoherent or completely irrelevant to the topic.

THINKING SKILLS

Paper 9694/22
Critical Thinking

General comments

Candidates seemed to respond well to the issues raised by the questions and were able to tackle them effectively, and there was a higher proportion who were well-prepared. The trend in recent papers of most candidates understanding the nature of the examination continued. However, there is still a minority of candidates who seem unprepared and struggle to reach a total mark in double figures.

Key messages

- Candidates should avoid over-long answers to 3-mark questions. In some cases this seems to have led to candidates being unable to complete **Question 3d**, thus losing a good opportunity to gain marks.
- It is not sufficient to simply identify *potentially* useful or relevant points (for example) from a source. Candidates must explain why they are useful/relevant.

Comments on specific questions

Question 1

- (a) Many candidates saw the tension between the pilot being an expert and an eye-witness but also having a vested interest in offering an explanation which exonerated him from blame. Many also pointed out how his account was possibly undermined by his statement in Source D. Some candidates tried to argue his statement was corroborated by Source A but nothing in Source A points to loss of power as such. This question was answered more successfully than similar questions in recent past papers.
- (b) Many candidates successfully argued that the expertise and neutrality of the source made it useful (usually the former). However they tended to just repeat points made in the passage (e.g. the plane was old) rather than explaining clearly why this information was useful (e.g. in offering a possible explanation for the incident). See the second key message above.
- (c) Candidates tackled this question less successfully. Only a (albeit significant) minority saw that the crucial question as regards relevance was whether the pilot in Source B was an older retired pilot. Some were confused and thought that the pilots in Source E had also retired from flying at air shows, leading to comments such as 'He is not retired because he is flying at the air show'.
- (d) There were usually clear conclusions, though some candidates seemed to change their mind, often finishing with a different conclusion than the one they stated at the beginning of their answer. Answers were fairly evenly split between arguments that the plane had lost power versus the idea that he had been performing risky crowd-pleasing stunts, probably with his bonus in mind. Most candidates rather assumed that the email did reveal collusion with the organisers and those candidates who questioned this were able to gain an evaluation of evidence mark. Some candidates with an interest in flying went in to rather technical speculative detail – it is important to stress one must confine one's answer to the evidence provided.

Question 2

- (a) This was quite an easy question and most candidates saw that a 'per capita' figure / population size would be needed in order to reliably conclude that the inhabitants had a lower level of income.

A minority of candidates saw this as a question about the reliability of the source, leading them to wrongly say it could be reliably concluded because it was an expert source.

- (b) This question was less well done, but a substantial number of candidates saw the key point about necessity of a car in the two different situations. However, very few candidates developed this point in relation to income i.e. the key thing that needed explaining was why the richer group had less car ownership than the poorer group. Some candidates confused population density with household size. Others simply described what the figures showed.
- (c) This question was done well, with many candidates gaining 3 marks by seeing that the definition of malnutrition in the source (lacking food that promoted healthy tissue and organ function) was compatible with being obese as a result of overeating food that failed to provide the nutrients etc. needed.
- (d) Whilst most candidates successfully disagreed with the claim by using Source A, many struggled to use the other sources to clearly tackle the question. Good answers made the point that even if inhabitants in high GDP countries had a reasonable share of that GDP as income, we would need to know cost of living figures in order to see whether that income was sufficient for them to avoid poverty. A minority of candidates were able to successfully use Source D to suggest that it raised definitional problems that GDP figures did not really tackle but others tended to simply repeat the points made in the source.

Question 3

- (a) Only a minority of candidates successfully identified the main conclusion, with the opening sentence being a popular but wrong answer.
- (b) In spite of not identifying the main conclusion, many candidates did succeed in identifying three intermediate conclusions. The vast majority of candidates identified at least one. In some cases the main conclusion was wrongly cited as an intermediate conclusion.
- (c) Many candidates found further compensation for failing to spot the main conclusion in their answers to this question. There were a large number gaining 3 marks or above with the generalisations, the confusion of ineffectiveness of alarms with ineffectiveness of response, the assumptions about immobilisers and the slippery slope in the final paragraph all being recognised. Some strong responses notionally exceeded the maximum 5 marks available. There is still, however, a minority of candidates who confuse evaluation of reasoning with critique of empirical basis and style of expression. Such candidates tend to berate the author for not presenting both sides of the argument.
- (d) Most candidates argued for the proposition. A fair number of responses interpreted 'noise' to mean all sound in general, hence arguing the importance of living with sound and citing music and birdsong to support their case. As this question is mainly intended to test whether candidates can use argument elements to create their own argument they were not unduly penalised for this approach. A few actively addressed the question of when a sound becomes a noise. A minority of candidates focused on alarms as such, often using material from the passage and this approach could not be credited much, if at all. There were some interesting arguments against the proposition, using health effects of noise effectively to make the case. Most candidates did take a consistent line either for or against, but a minority of candidates treated the exercise more as a review of the issue, looking at both sides, and this was a less successful approach, given the nature of the question.

THINKING SKILLS

Paper 9694/23
Critical Thinking

Key messages

Teachers and students 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 give a correct answer.

The number of marks allocated to a question is a good guide to the complexity and length of answer expected.

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.

General comments

Most candidates had time to attempt all questions, and with a few exceptions the omissions were probably either accidental or because candidates did not know how to answer the question.

Comments on specific questions

Question 1

Nearly all candidates understood the issue and the implications of the sources.

- (a) Nearly all candidates recognised that the additional information reduced the reliability of the report, but many comments were too undeveloped to achieve the full three marks available. Some candidates stated that the additional information 'affected' the reliability (which the question had already implied) but omitted to state whether it increased or reduced the reliability, thereby losing the judgement mark. A few candidates summarised or discussed the content of the report instead of evaluating its reliability.
- (b) Candidates appear to have found this question difficult, and only a few achieved three marks. Although the wording of the question implied that the statement was significant and did not ask for a judgement, some candidates replied that it was not significant.
- (c) Many candidates succeeded in drawing relevant inferences from this event in relation to Mr Horak's presidency, but fewer recognised that the new government's evident anti-Eastish bias made it more likely that the charges against Mr Jalin were motivated by revenge or sectarian bias. Although the wording of the question implied that the additional information was significant and did not ask for a judgement, some candidates replied that it was not significant. Some candidates implausibly thought that the religious building had been demolished while Marcel Jalin was President, while others accepted the claim that the building was unsafe; both of these approaches made it hard for candidates to make pertinent comments.

- (d) There were some very good answers to this question, some of them exceeding the requirements for full marks. Many candidates made good use of inferential reasoning from the sources. Many candidates judged that the accusations against Mr Jalin were false, but it was also possible to argue well for the opposite point of view. Quite a lot of candidates thought that if soldiers killed protestors, this would make the President guilty of murder. Some candidates gave good discussions of an alternative to their own judgement.

Question 2

- (a) Most candidates correctly judged that the evidence was reliable, although some thought the long timescale of the evidence studied was a weakness rather than a strength. Some candidates suggested that Prof Phillips did not have the appropriate expertise, because he is a Professor of Sociology rather than Medicine, but this was not credited, because research into patterns of mortality does not require medical knowledge. A few candidates summarised or discussed the content of the research report instead of evaluating its reliability.
- (b) Some weak answers discussed the credibility of the nurse rather than comparing the content of her comments with those in Source A, as the question intended. Many candidates recognised that the conclusion of the two documents were similar, and jumped to the incorrect conclusion that Source B supported Source A. A significant minority of candidates performed well on this question, by comparing the two documents carefully and correctly identifying the points of connection and those of disconnection.
- (c) Some candidates understood the expression 'unstated assumption' and recognised that the Health Minister was relying on the assumption that the seasonal increase in deaths was preventable, but the most popular answer was quotation or paraphrase of the second sentence from the passage, which is by definition not 'unstated'. Many other incorrect answers were also given.
- (d) As on previous occasions, many candidates accurately summarised the sources and drew an acceptable conclusion, thereby achieving three marks out of six, while others accessed the higher marks, by making good use of evaluation and/or inferential reasoning.

Question 3

This topic was slightly more abstract and intellectual than some, but most candidates engaged with it well.

- (a) A fair proportion of candidates correctly identified the main conclusion, even though it was not located in a prominent position. Many candidates gave wrong answers, the most popular of which 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. The final sentence of paragraph 1 was a popular wrong answer.
- (c) A fair proportion of candidates understood what was expected by this question, and gave at least one correct answer. Popular correct answers were the false dichotomy in paragraph 2, the straw man (or slippery slope) in paragraph 4 and the *ad hominem* in paragraph 5. Some candidates wrongly criticised the first sentence for being a generalisation or an appeal to tradition (probably because the word 'tradition' occurs in that sentence).

As on previous occasions, a significant proportion of candidates misunderstood the nature of this task, explaining why they agreed or disagreed with all or some of the opinions and claims in the passage, instead of evaluating the reasoning. Evaluative points which were not credited included criticism of the argument for being one-sided, for expressing the opinion of the implied author, or for lacking statistical support. Literary criticism was also not credited.

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

- (d) Many candidates responded well to this question, most choosing to support the claim. Candidates who opposed the claim tended to interpret it as meaning that one should benefit others in preference to oneself, which is not what was stated or intended. Many made good use of appropriate examples and organised their arguments in strands of reasoning. However, some discussed both sides of the issue without supporting one side or the other, thereby not doing what the question asked them to do, and were therefore awarded only one mark (for a relevant comment). A few incoherent or irrelevant answers were awarded no marks. A few candidates misinterpreted the claim as meaning that people should do their best (at everything they do), because this will benefit others.



THINKING SKILLS

Paper 9694/31
Problem Analysis and Solution

Key messages

Under the time pressure of the exam, candidates may benefit from spending a few minutes glancing through the questions before deciding which one to leave until last.

Although in some questions it may be helpful to do some working on the question paper, candidates should transcribe key aspects of this working to their answer script to ensure that they have a chance of scoring partial credit where it is available.

General comments

Candidates found **Question 2** the most difficult. Most candidates took sensible approaches to **Questions 1, 3 and 4**, and many left comprehensible working, which allowed for the award of partial credit. Some candidates appeared to run out of time, leaving parts of **Question 4** unfinished or cursorily answered see the second key message above.

Comments on specific questions

Question 1

This question required candidates to consider the overlapping intervals determined by different sequences of numbers – and how this affected the ‘long term cost’ of taking pills on those days. The majority of candidates made successful attempts at the first four parts – which considered only the frequency of the pills and the intervals that remained – and struggled with the rates required in calculating a long term cost.

- (a) Many candidates listed the dates in a calendar format, deleted the dates from the relevant sequences, and accomplished the question successfully. The most common misconception seen in candidates’ answers was to assume that an easy shortcut was available when considering the number of dates in each of the two lists separately.
- (b) This question added a second restriction to Fred’s social habits, which could be overlaid onto the calendar created for **part (a)**. A number of candidates did this successfully. Many also just gave the number, without explanation, which was awarded full marks. The most common error was to conclude ‘10 days’, without explanation.
- (c) This question took the familiarity that candidates had gained with the sequences, and posed a limited ‘optimisation’ problem based on this. This immediately invoked a need for strategic thinking, since a consideration of all the dates in June was not recommended (given the pressure on time). Many candidates realised that the 5th was the last day that the pill could be taken (if he is to take six pills as required). Some attempted to encode the possibilities on the (cluttered) diagram from **part (a)** – which yielded mixed results. Those who offered some rationale for their answers tended to gain 1 mark even if they made computational errors.
- (d) This question was a straightforward processing question, although the instructions were more complex than those encountered so far. Some candidates drew out basic calendar grids for each of the new pills – and many candidates offered the correct pair of results.
- (e) This question involved a different perspective on the sequences, asking candidates to consider the ‘long-term’ cost. This could be tackled in a variety of ways, most commonly by taking an arbitrarily large number of days (often candidates used 400, although this was not sufficiently large to

determine the long term trend), or by trying to find a price per day. This was accomplished by far fewer candidates. Many candidates calculated and compared the price per box – but this was not credited unless the frequency and number of pills in each box were shown. This calculation also required careful handling and many candidates selected the wrong figure when comparing the expiry date and the potential life of a box (calculated as the frequency multiplied by the number of pills, or number – 1).

Question 2

This question required candidates to investigate how a collection of proportions, iterations and absolute values interacted in a model representing the number of cars outside a school. The relationship between the structural proportions of the problem and the snapshots of information about the number of cars parked had to be sensitively applied, and many candidates did not manage to fully master the mechanism and the questions that arose.

- (a) This question involved a simple iteration – a constant addition of students, bringing with it 10 cars, for 5 years – but the problem was overshadowed by the fact that the proportion of students who were Nesters and Cuckoos was not given. It can be surmised that those who realised that this proportion was not necessary found this question easy. Those who were distracted by it may have found themselves tackling more than half the question (worth six marks) before properly beginning.
- (b)(i) This simple question involved separating two pieces of information out from the model: 20 Cuckoos were added in the first year; the number of cars increased by 10. As with **part (a)**, there was the distracting absence of information about the number of Nesters – and the pursuit of this often led to the hasty adoption of inappropriate assumptions; for instance, a number of candidates assumed that half of the children were Cuckoos, and none of the Nesters travelled by car, yielding a proportion of 35/50 or 70 percent.
- (ii) Those candidates who gave a wrong answer to **(b)(i)** were unlikely to gain a mark here, because the follow-through calculations tended to give impossible answers. The appropriate logic required the application of the reciprocal proportion from **(b)(i)** to the number of cars: 50 percent of Cuckoos travel by car, implies that there are twice as many Cuckoos as cars, which implies 70 Cuckoos in the first year. That leaves 30 Nesters.
- (c) This question finally dealt with the piece of information which had been so distractingly absent in the first three questions: how many Nesters are at the school? Candidates were awarded partial marks for pursuing the correct approach here, even if some of their previous answers were incorrect: a comparison of the two numbers of cars after the first year (55 if the numbers increased as the first year; 52 if all the final year Nesters stay at home), a deduction of how many cars from Nesters there were in the whole school (five years, so five times three), and then an inference to how many cars from Cuckoos there were originally, and finally a conclusion with regard to the number of Nesters. Quite a few candidates managed the first step or two of this; very few reached the end.
- (d) With all the basic information in hand, this question simply involved iterating the process carefully: 40 Cuckoos and 60 Nesters, with proportions of 50 percent and 25 percent coming in cars respectively, with four years of students parking, gives $(3 + 14) + (3 + 14) + (3 + 14) + (3 + 4) = 58$ cars. There were a number of alternative ways of dividing up the numbers; very few candidates left working which was sufficiently clear to attract partial credit (if the final answer was incorrect).
- (e) A similar question to **(d)**, which could be tackled straightforwardly given the correct proportions. Many candidates had made inadvertent errors by this point, and follow through marks were available – but a clear method had to be visible to attract partial credit.

Question 3

Question 3 followed a week in the life of two tailors – and how they blocked their time making trousers, waistcoats and jackets. The familiarity of this situation ensured that a high number of candidates attempted most of the question. And many candidates were awarded marks for their working, even if leading to an incorrect conclusion, since it did not require much supporting text to make the working unambiguous.

- (a) This question checked candidates understanding of the prices and the discount, and was completed well by most candidates.

- (b) A simple diagram or model of the time available to the two tailors enabled candidates to forecast the completion for Thursday. There was no need to consider the possible combinations here, since it was clear that one would make the trousers, and the other would make the jacket, each taking 20 hours = 8 + 8 + 4. The information in the preceding paragraph did need to be extracted and simplified in order to be processed – and about one quarter of the candidates did not do this.
- (c) This question moved the task up to a tactical level – with some consideration needed of which items earned the tailors the most money, and how the discount might interfere with this. Most candidates selected one of the three combinations which could be completed in 30 hours, but many did not then check that they had found a maximum. The most common error was to forget the discount (and claim that \$250 could be made)
- (d) This question required some consideration of the combinations available, and was most effectively done by those who invoked a diagram or tabular structure of some sort. At a strategic level, most candidates appreciated that finding the total amount of time and halving it established a goal for minimising their overtime. There were a number of ways that the items could be split into 115 and 120 hours.
- (e) This question was similar in structure to (d), but requiring more care as a result of the tailors' unequal starting points. This question was harder to accomplish without some diagramming or tabular representation, and it was completed less successfully by many candidates.
- (f) A correct solution to this required appreciation that the items must be arranged to take up as much of Harry's time as possible. It is clear how to occupy 35 of his remaining 36 hours, and thus the answer is one less than the answer to (e). Follow through marks were available here, but few candidates managed the task correctly.
- (g) This question could be tackled separately from the previous scheduling problems, and required a careful consideration of how much the two non-urgent orders would be reduced if not completed until Monday: a \$40 order which was three days late, compared to a \$90 order which was one day late. Only a few candidates managed this correctly.

Question 4

This question offered a collection of problems based around the sales and bonuses that could be awarded to a group of employees. The problems emerged from the relationships between the sales, bonuses, and points – none of the relationships between these were complex.

- (a) This question tested candidates' ability to 'unpack' the table. No working was necessary to gain the mark. Most candidates completed this correctly.
- (b) Success at this question depended on thoroughness and care in the award of the monthly bonuses. Some working was likely to have occurred on the exam paper, but this does not reach the eyes of the examiner. Any working which candidates wish to leave for public consumption (i.e. anything which could gain marks, or anything which shows the choices that have been made, in a multi-mark question) must be written in the answer booklet.
- (c) A table demonstrating how sales are converted to points was introduced and interrogated in this question. As with all tables, it paid to check that how the cells are to be interpreted. In particular, it is tempting to conclude from a table of this kind that the points per sale (in the cells) can be multiplied by the number of sales (column headings) to give the number of points, e.g. 21 sales in the East quadrant seems to be worth $21 \times 2 = 42$ points. But the example under the table shows that this is not correct. The most common incorrect answer to this question was '2'.
- (d) Those who interpreted the point/sales table too hastily were severely restricted in this question. Many candidates may have left their working on their question paper – but the four marks available and the detail required in the task made it advisable to leave evidence for the examiner. Those who erroneously included the monthly bonus system (from the beginning of the question) were able to obtain partial credit.

- (e) This application of the new rule was no different in structure to the six-month rule, and the focus on Martin made the task easier. Most candidates did not offer the two figures required, and this may have reflected the time pressures.
- (f) This final question required handling the inverse of the stepped relationship between bonuses, points and sales. Any proposed answers could be verified fairly easily (using the table, operated the familiar way round), but this is often hard to manage on the final question of a time-bound paper.

THINKING SKILLS

<p>Paper 9694/32 Problem Analysis and Solution</p>
--

Key messages

As the title suggests, this component is concerned with both analysis and solution, not just the final answer. Candidates need to be able to explain and justify their answers, particularly when this is the focus of the question.

General comments

Candidates need to distinguish between part questions which develop the scenario in different ways and those which offer steps towards an ultimate answer. When the scenario is amended, care is needed to be aware whether earlier work is still relevant.

Questions where four marks are available can be expected to have at least four steps for a complete answer, whether these are explanations or terms in a calculation.

This paper did not involve any arithmetic with times, so there were fewer arithmetic errors than previously observed.

Comments on specific questions

Question 1

This question involved considering various limit cases in a contrived but simple model. Some candidates added constraints of their own.

- (a) (i) This part required considering one extreme, and determining how many groups could be affected by a change. A few interpreted it as a generic question and ignored the information immediately preceding the question.
- (ii) The opposite limit was explored here. In this context candidates should consider that the smallest number is zero, not one.
- (b) This part required calculation of the minimum possible points for each group, added up and then converted into dollars.
- (c) This merely sought identification that the amount available was less than the minimum needed. Many candidates introduced a constraint not in the question: a limit of one tutor per student.
- (d) For full marks it was necessary to include both those who already made the grade and those whose mark could be boosted. Many only considered the latter.

Question 2

This question explored which colours were assigned to visitor badges on different days, using a mixture of observations and an outline description of the method.

- (a) Most candidates determined that four would be needed. Where working was shown there were different acceptable methods: more than three and four works, or it must be a divisor of $(18 - 2) = 16$ greater than 2.
- (b) Many candidates assumed that George thought that the smallest number (as determined in (a)) was used.
- (c) (i) For full marks some working was required to show that the response is not a guess but a logical deduction from the information given. This includes noting that it must be the answer rather than simply could be.
(ii) Different but equally creditworthy approaches were adopted: finding specific contradictions or studying the possible pattern lengths for lack of clashes.
- (d) The day(s) for the missing colour could be found either looking at what was left by filling the known colour days or by looking at the pattern of the missing colour days during the first four weeks.

Question 3

This question involved a model for scheduling the construction of a railway, using up rock from cuttings to make embankments and disposing of the surplus.

- (a) Most candidates saw that this part required nothing more than adding up the numbers, both positive and negative.
- (b) (i) Most candidates observed that, starting from the left, there is not enough rock to fill section QP.
(ii) The majority noted that there were enough blocks on the right to fill both stretches where embankments were needed.
- (c) (i) This asked for a point. Most candidates offered a segment. Some neglected the additional constraint about temporary storage and re-used the answer to (b)(ii).
(ii) Few candidates noticed features in the notation designed to reduce errors, e.g. the underlining to show where tracks lay. Some did not observe that the notation indicated the position *after* each move.
- (d) The calculation required consideration of all the blocks moved, included those discarded.
- (e) There are many ways to obtain the cheapest cost; this part required working out what guarantees cheapest, and selecting an example, (preferably written using the notation provided). Some candidates offered a cheapest that was, implausibly, more than their answer to (d).
- (f) The reduction in cost would be from two considerations: not needing to move the excess rock so far to dispose of it and the price for which it was sold. Many determined the latter value of four blocks at \$350 each, but few noted the former. The allocation of two marks might have been used as a hint that something more complicated than multiplying two given numbers was expected.

Question 4

This question involved analysis of the results of a usual three-way game, as well as some exploration of the future scores needed by the teams.

- (a) (i) Most candidates correctly noted the first starting time of the last of the eight teams.
(ii) The first team to complete their matches could be obtained either by working backwards or by establishing the number of appearances and working forwards.

- (b) (i)** Full marks required an explanation of how some of the points *must* have been awarded.
- (ii)** Most candidates established the missing scores, but some did not present them in the format specified.
- (c) (i)** Although not difficult to note that $6 + 2 =$ (all) 8 (goals), some candidates omitted the explanation required for the mark.
- (ii)** The number of people, not the number of goals was wanted. The identity of those who only scored one goal is not given, but they still needed to be included in the count.
- (iii)** All potential winners were wanted. Many candidates overlooked constraints such as the score of a player cannot exceed that of his team.
- (d) (i)** The analysis of what happens if no *further* goals are scored requires consideration of all goals so far; most are given in a table but some are mentioned in the text.
- (ii)** This part required an explanation as to why the given number of goals was sufficient to reach the final. It opened up the exploration needed for the more complex next part.
- (iii)** Here it asked for an explanation of how a specific team reach the final, but only under a particular condition. Few candidates addressed the necessity of the 'only if' condition, although some used it as a hint.

THINKING SKILLS

Paper 9694/33
Problem Analysis and Solution

Key messages

Under the time pressure of the exam, candidates may benefit from spending a few minutes glancing through the questions before deciding which one to leave until last.

Although in some questions it may be helpful to do some working on the question paper, candidates should transcribe key aspects of this working to their answer script to ensure that they have a chance of scoring partial credit where it is available.

General comments

Candidates found **Question 2** the most difficult. Most candidates took sensible approaches to **Questions 1, 3 and 4**, and many left comprehensible working, which allowed for the award of partial credit. Some candidates appeared to run out of time, leaving parts of **Question 4** unfinished or cursorily answered see the second key message above.

Comments on specific questions

Question 1

This question required candidates to consider the overlapping intervals determined by different sequences of numbers – and how this affected the ‘long term cost’ of taking pills on those days. The majority of candidates made successful attempts at the first four parts – which considered only the frequency of the pills and the intervals that remained – and struggled with the rates required in calculating a long term cost.

- (a) Many candidates listed the dates in a calendar format, deleted the dates from the relevant sequences, and accomplished the question successfully. The most common misconception seen in candidates’ answers was to assume that an easy shortcut was available when considering the number of dates in each of the two lists separately.
- (b) This question added a second restriction to Fred’s social habits, which could be overlaid onto the calendar created for **part (a)**. A number of candidates did this successfully. Many also just gave the number, without explanation, which was awarded full marks. The most common error was to conclude ‘10 days’, without explanation.
- (c) This question took the familiarity that candidates had gained with the sequences, and posed a limited ‘optimisation’ problem based on this. This immediately invoked a need for strategic thinking, since a consideration of all the dates in June was not recommended (given the pressure on time). Many candidates realised that the 5th was the last day that the pill could be taken (if he is to take six pills as required). Some attempted to encode the possibilities on the (cluttered) diagram from **part (a)** – which yielded mixed results. Those who offered some rationale for their answers tended to gain 1 mark even if they made computational errors.
- (d) This question was a straightforward processing question, although the instructions were more complex than those encountered so far. Some candidates drew out basic calendar grids for each of the new pills – and many candidates offered the correct pair of results.
- (e) This question involved a different perspective on the sequences, asking candidates to consider the ‘long-term’ cost. This could be tackled in a variety of ways, most commonly by taking an arbitrarily large number of days (often candidates used 400, although this was not sufficiently large to

determine the long term trend), or by trying to find a price per day. This was accomplished by far fewer candidates. Many candidates calculated and compared the price per box – but this was not credited unless the frequency and number of pills in each box were shown. This calculation also required careful handling and many candidates selected the wrong figure when comparing the expiry date and the potential life of a box (calculated as the frequency multiplied by the number of pills, or number – 1).

Question 2

This question required candidates to investigate how a collection of proportions, iterations and absolute values interacted in a model representing the number of cars outside a school. The relationship between the structural proportions of the problem and the snapshots of information about the number of cars parked had to be sensitively applied, and many candidates did not manage to fully master the mechanism and the questions that arose.

- (a) This question involved a simple iteration – a constant addition of students, bringing with it 10 cars, for 5 years – but the problem was overshadowed by the fact that the proportion of students who were Nesters and Cuckoos was not given. It can be surmised that those who realised that this proportion was not necessary found this question easy. Those who were distracted by it may have found themselves tackling more than half the question (worth six marks) before properly beginning.
- (b)(i) This simple question involved separating two pieces of information out from the model: 20 Cuckoos were added in the first year; the number of cars increased by 10. As with **part (a)**, there was the distracting absence of information about the number of Nesters – and the pursuit of this often led to the hasty adoption of inappropriate assumptions; for instance, a number of candidates assumed that half of the children were Cuckoos, and none of the Nesters travelled by car, yielding a proportion of 35/50 or 70 percent.
- (ii) Those candidates who gave a wrong answer to **(b)(i)** were unlikely to gain a mark here, because the follow-through calculations tended to give impossible answers. The appropriate logic required the application of the reciprocal proportion from **(b)(i)** to the number of cars: 50 percent of Cuckoos travel by car, implies that there are twice as many Cuckoos as cars, which implies 70 Cuckoos in the first year. That leaves 30 Nesters.
- (c) This question finally dealt with the piece of information which had been so distractingly absent in the first three questions: how many Nesters are at the school? Candidates were awarded partial marks for pursuing the correct approach here, even if some of their previous answers were incorrect: a comparison of the two numbers of cars after the first year (55 if the numbers increased as the first year; 52 if all the final year Nesters stay at home), a deduction of how many cars from Nesters there were in the whole school (five years, so five times three), and then an inference to how many cars from Cuckoos there were originally, and finally a conclusion with regard to the number of Nesters. Quite a few candidates managed the first step or two of this; very few reached the end.
- (d) With all the basic information in hand, this question simply involved iterating the process carefully: 40 Cuckoos and 60 Nesters, with proportions of 50 percent and 25 percent coming in cars respectively, with four years of students parking, gives $(3 + 14) + (3 + 14) + (3 + 14) + (3 + 4) = 58$ cars. There were a number of alternative ways of dividing up the numbers; very few candidates left working which was sufficiently clear to attract partial credit (if the final answer was incorrect).
- (e) A similar question to **(d)**, which could be tackled straightforwardly given the correct proportions. Many candidates had made inadvertent errors by this point, and follow through marks were available – but a clear method had to be visible to attract partial credit.

Question 3

Question 3 followed a week in the life of two tailors – and how they blocked their time making trousers, waistcoats and jackets. The familiarity of this situation ensured that a high number of candidates attempted most of the question. And many candidates were awarded marks for their working, even if leading to an incorrect conclusion, since it did not require much supporting text to make the working unambiguous.

- (a) This question checked candidates understanding of the prices and the discount, and was completed well by most candidates.

- (b) A simple diagram or model of the time available to the two tailors enabled candidates to forecast the completion for Thursday. There was no need to consider the possible combinations here, since it was clear that one would make the trousers, and the other would make the jacket, each taking 20 hours = $8 + 8 + 4$. The information in the preceding paragraph did need to be extracted and simplified in order to be processed – and about one quarter of the candidates did not do this.
- (c) This question moved the task up to a tactical level – with some consideration needed of which items earned the tailors the most money, and how the discount might interfere with this. Most candidates selected one of the three combinations which could be completed in 30 hours, but many did not then check that they had found a maximum. The most common error was to forget the discount (and claim that \$250 could be made)
- (d) This question required some consideration of the combinations available, and was most effectively done by those who invoked a diagram or tabular structure of some sort. At a strategic level, most candidates appreciated that finding the total amount of time and halving it established a goal for minimising their overtime. There were a number of ways that the items could be split into 115 and 120 hours.
- (e) This question was similar in structure to (d), but requiring more care as a result of the tailors' unequal starting points. This question was harder to accomplish without some diagramming or tabular representation, and it was completed less successfully by many candidates.
- (f) A correct solution to this required appreciation that the items must be arranged to take up as much of Harry's time as possible. It is clear how to occupy 35 of his remaining 36 hours, and thus the answer is one less than the answer to (e). Follow through marks were available here, but few candidates managed the task correctly.
- (g) This question could be tackled separately from the previous scheduling problems, and required a careful consideration of how much the two non-urgent orders would be reduced if not completed until Monday: a \$40 order which was three days late, compared to a \$90 order which was one day late. Only a few candidates managed this correctly.

Question 4

This question offered a collection of problems based around the sales and bonuses that could be awarded to a group of employees. The problems emerged from the relationships between the sales, bonuses, and points – none of the relationships between these were complex.

- (a) This question tested candidates' ability to 'unpack' the table. No working was necessary to gain the mark. Most candidates completed this correctly.
- (b) Success at this question depended on thoroughness and care in the award of the monthly bonuses. Some working was likely to have occurred on the exam paper, but this does not reach the eyes of the examiner. Any working which candidates wish to leave for public consumption (i.e. anything which could gain marks, or anything which shows the choices that have been made, in a multi-mark question) must be written in the answer booklet.
- (c) A table demonstrating how sales are converted to points was introduced and interrogated in this question. As with all tables, it paid to check that how the cells are to be interpreted. In particular, it is tempting to conclude from a table of this kind that the points per sale (in the cells) can be multiplied by the number of sales (column headings) to give the number of points, e.g. 21 sales in the East quadrant seems to be worth $21 \times 2 = 42$ points. But the example under the table shows that this is not correct. The most common incorrect answer to this question was '2'.
- (d) Those who interpreted the point/sales table too hastily were severely restricted in this question. Many candidates may have left their working on their question paper – but the four marks available and the detail required in the task made it advisable to leave evidence for the examiner. Those who erroneously included the monthly bonus system (from the beginning of the question) were able to obtain partial credit.

- (e) This application of the new rule was no different in structure to the six-month rule, and the focus on Martin made the task easier. Most candidates did not offer the two figures required, and this may have reflected the time pressures.
- (f) This final question required handling the inverse of the stepped relationship between bonuses, points and sales. Any proposed answers could be verified fairly easily (using the table, operated the familiar way round), but this is often hard to manage on the final question of a time-bound paper.

THINKING SKILLS

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

Key messages

- In **Question 1** many candidates were able to describe several criticisms of the study and most gained some marks.
- In **Question 3** most candidates did as they were asked and attempted to evaluate the reasoning.
- In **Question 4**, candidates can only achieve the highest marks if create their own argument structure, ignoring the sequence in which the documents are presented, and if they engage *critically* with the documents provided.

General comments

Most candidates appeared to have enough time to finish the paper but a few did not. Often, those who did not have time to complete **Question 4** had spent a disproportionately long time on previous questions, although such responses were in the minority.

The standard of candidates varied but there was evidence that many candidates had been well prepared with regard to answering **Questions 1, 2 and 3**.

Comments on specific questions

Question 1

Candidates scored better on this question than in many previous **Question 1s**. The full range of scores, from 0 to 5, was seen. Most commonly candidates gained marks for reference to visitor non-compliance with advice and the potential atypicality of the hospital or the year of study. A smaller number addressed the lack of comparison with the other 15 wards, the significance of the 8% or the misleading rounding in the headline figure. Surprisingly few expanded on the atypicality idea with reference to the idea that the hospital was a centre of excellence for major surgery. It was encouraging to see very few candidates criticise the credibility of the authors of the report.

Question 2

This question rewarded the well-prepared candidate. Those who knew what was expected and attempted an analysis of the argument usually gained five marks easily. Although there were fewer elements to identify, candidates were generally able to identify the ones that were there. Very few candidates provided a non-creditworthy summary or gist. As ever, some candidates are still unaware that quoting from the text is an appropriate, indeed a required, way to answer this question.

Question 3

Most candidates were aware of the nature of the task and attempted to evaluate the passage. Those candidates were often able to gain some marks fairly easily. The most frequently credited weaknesses identified were the contradiction between the claim that newly-discovered animals are small and the example of the colossal squid. All other weakness points on the mark scheme were seen and credited. Historically, very few candidates are able to identify assumptions, indeed many appear not to have learned that an assumption must not be stated in the text. However, on this occasion, all the assumption points, with the exception of the first in paragraph 3, were seen and 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 the yeti is nothing more than a myth. Many found they could engage with this topic, and more candidates than usual seemed particularly ready and able to offer appropriate counter-arguments to points brought up in the documents. More marks than usual were awarded for the 'treatment of counter-positions' skill. It was noted in the June 2018 session that many candidates were able to arrange their ideas into strands of reasoning with very little sequential treatment of documents. However, in the current series many candidates presented their answers as a series of analyses of each document in turn. Hence, marks for the structure and quality skills were, for some candidates, unable to rise above Level 1 or 2. Some candidates were using the documents with a critical eye, which allowed their marks for 'use of documents' to access Level 3 or above. It is worth reminding centres that 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 relevant alternative viewpoints.

THINKING SKILLS

Paper 9694/42
Applied Reasoning

Key messages

- In **Question 1** the quality of responses varied greatly, but some were able to gain marks.
- In **Question 3** some candidates are still gaining 0 marks because they are 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, if not all, candidates appeared to have enough time to finish the paper. Some who appeared to rush **Question 4** had spent a disproportionately long time on previous questions, although such responses were in the minority.

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

Comments on specific questions

Question 1

Candidates were asked to criticise the study outlined in the report or any inferences drawn from it. The quality of responses varied between 0 and 4 marks, with most candidates achieving 1 or 2 marks. Most commonly, credit was given for reference to potential confounding variables, suggestion that the baseline ability of students might have changed or, more rarely, that the 2.2% might be within normal annual fluctuations, and a few alluded to problems with ambiguities surrounding intelligence, exam success and brain development. Occasionally candidates were able to expand on some of these points for extra credit, referencing the absence of a control group or effects related to the very existence of a study. It was good that fewer candidates than usual devoted a large proportion of their answer to questioning the credibility or provenance of the data presented.

Question 2

As ever, this question rewarded the well-prepared candidate. Those who knew what was expected and attempted an analysis of the argument usually gained four marks easily. The main conclusion was relatively easy to spot so no responses were capped at 4 marks. A few candidates provided a non-creditworthy summary or gist. As ever, some candidates are unaware that quoting from the text is an appropriate, indeed a required, way to answer this question. A very few attempted to evaluate the reasoning, like 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. Those candidates who did attempt to apply their evaluation skills were often able to gain some marks easily. Despite the small entry, almost all of the points on the mark scheme were seen; some gaining full credit, and some given 1 mark if they were associated with a lack of clarity or use of the wrong technical term. Marks ranged between 0 and 8. No candidate suggested that the statement “We do not like the uniform because it makes us unhappy” was circular. Historically, very few

candidates are able to identify assumptions, indeed many appear not to have learned that an assumption must not be stated in the text. On this occasion this misunderstanding was also in evidence, but there were some candidates who were able to correctly identify some unstated assumptions.

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 all schools should have compulsory school uniform. Most candidates stated a clear conclusion and some were able to organise their ideas into strands of reasoning. However, there was little use of explicit intermediate conclusions and other argument elements. Hence no candidate achieved Level 4 in the 'structure' skill. However, most candidates did attempt to structure their own arguments, rather than merely appraising the documents sequentially, so most candidates scored above Level 1 for this skill. There was evidence that candidates were engaging with the topic but the manifestation of this was that candidates frequently offered counter assertions to issues raised in the documents rather than positive ideas of their own. This meant that marks were awarded for treatment of counter positions, more often than usual, but the 'quality of argument' marks were similarly a little lower than usual. Responses often contained examples of the candidates own thinking and ideas. However, most candidates are still using the documents without a critical eye, which limits their marks for use of documents but also, necessarily, detracts from the persuasiveness of their case. It is worth reminding Centres that 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 relevant alternative viewpoints.

THINKING SKILLS

Paper 9694/43
Applied Reasoning

Key messages

- In **Question 1** many candidates were able to describe several criticisms of the study and most gained some marks.
- In **Question 3** most candidates did as they were asked and attempted to evaluate the reasoning.
- In **Question 4**, candidates can only achieve the highest marks if create their own argument structure, ignoring the sequence in which the documents are presented, and if they engage *critically* with the documents provided.

General comments

Most candidates appeared to have enough time to finish the paper but a few did not. Often, those who did not have time to complete **Question 4** had spent a disproportionately long time on previous questions, although such responses were in the minority.

The standard of candidates varied but there was evidence that many candidates had been well prepared with regard to answering **Questions 1, 2 and 3**.

Comments on specific questions

Question 1

Candidates scored better on this question than in many previous **Question 1s**. The full range of scores, from 0 to 5, was seen. Most commonly candidates gained marks for reference to visitor non-compliance with advice and the potential atypicality of the hospital or the year of study. A smaller number addressed the lack of comparison with the other 15 wards, the significance of the 8% or the misleading rounding in the headline figure. Surprisingly few expanded on the atypicality idea with reference to the idea that the hospital was a centre of excellence for major surgery. It was encouraging to see very few candidates criticise the credibility of the authors of the report.

Question 2

This question rewarded the well-prepared candidate. Those who knew what was expected and attempted an analysis of the argument usually gained five marks easily. Although there were fewer elements to identify, candidates were generally able to identify the ones that were there. Very few candidates provided a non-creditworthy summary or gist. As ever, some candidates are still unaware that quoting from the text is an appropriate, indeed a required, way to answer this question.

Question 3

Most candidates were aware of the nature of the task and attempted to evaluate the passage. Those candidates were often able to gain some marks fairly easily. The most frequently credited weaknesses identified were the contradiction between the claim that newly-discovered animals are small and the example of the colossal squid. All other weakness points on the mark scheme were seen and credited. Historically, very few candidates are able to identify assumptions, indeed many appear not to have learned that an assumption must not be stated in the text. However, on this occasion, all the assumption points, with the exception of the first in paragraph 3, were seen and 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 the yeti is nothing more than a myth. Many found they could engage with this topic, and more candidates than usual seemed particularly ready and able to offer appropriate counter-arguments to points brought up in the documents. More marks than usual were awarded for the 'treatment of counter-positions' skill. It was noted in the June 2018 session that many candidates were able to arrange their ideas into strands of reasoning with very little sequential treatment of documents. However, in the current series many candidates presented their answers as a series of analyses of each document in turn. Hence, marks for the structure and quality skills were, for some candidates, unable to rise above Level 1 or 2. Some candidates were using the documents with a critical eye, which allowed their marks for 'use of documents' to access Level 3 or above. It is worth reminding centres that 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 relevant alternative viewpoints.

