

# THINKING SKILLS

---

**Paper 9694/21**  
**Critical Thinking**

## Key messages

- Candidates and teachers should study previous question papers and mark schemes, in order to familiarise themselves with the kinds of answers expected to the various questions.
- Candidates and teachers should familiarise themselves with the expectations attached to the prompt words used in questions, such as the difference between “reliable” and “significant” in Q1 and between “How reliable” (in Q1) and “Can it be reliably concluded” (in Q2).
- Candidates need to keep a clear head when evaluating evidence and arguments and avoid thinking that the fact that they disagree with a conclusion constitutes a valid criticism of the evidence or argument.

## General comments

A few candidates performed well overall, but there were a lot of mediocre or weak answer papers. The achievement of many candidates seems to have been compromised by their lack of sympathy for the views expressed in **Questions 2** and **3**.

## Comments on specific questions

### **Question 1**

Virtually all candidates seemed to find this scenario accessible and credible.

- (a) Nearly all candidates scored at least 1 mark on this question. Most recognised that the prohibition on demolishing Litho Hall was significant in some way, and some explained that it provided a motive for arson.
- (b) Most candidates judged correctly that Councillor Fox’s evidence was unreliable. Many of them supported this judgment by reference to the lack of corroborative evidence. Candidates who were familiar with credibility criteria tended to focus on the key point, his vested interest to blame other people in order to protect himself from suspicion.
- (c) As indicated on the mark scheme, several significant points could be made in response to this question, and most candidates identified at least one of them. Some discussed reliability instead of significance, while others discussed the evidence of the Chief Fire Officer instead of the police spokesman.
- (d) Nearly all candidates achieved a mark of at least 2, but above that there was a wide range of marks. Most candidates blamed Councillor Fox for the fire, but many of them expressed their judgment more confidently than the evidence supported; as the levels grid on the mark scheme makes clear, valid judgments in these questions are always a matter of probability rather than certainty. Some candidates produced good inferential reasoning in their use of evidence.

### **Question 2**

Many candidates performed very weakly on this question, apparently because they did not approve of the findings of the research.

- (a) Large numbers of candidates did not realize that the prompt words “Can it be reliably concluded” are intended to test the ability to draw an inference from evidence (Learning Objective CT10). The

point of this question was whether the findings of the research concerning School-age children could be extrapolated to university candidates, but many candidates instead answered it by criticizing the research report itself, on the grounds that it was biased, unsupported by data or unfair. These answers were not credited. Very few candidates explained that because entry to university is competitive, all candidates compete on equal terms once they have been admitted, but a fair number pointed out that differences in maturity are largely eroded by the time candidates enter university.

- (b) As explained in a preamble to this question, the point is that in adulthood, sports players compete against people from different year-groups, and it should therefore make no difference at what point in the year they were born, but in practice the Birthdate Effect is apparent in relation to professional players. Candidates were expected to explain how differences in Schooldays could have knock-on effects in adulthood. As indicated on the mark scheme, there were several valid answers, and a fair number of candidates offered one or two of them. Despite the explanatory introduction, however, many candidates assumed that a few months' age-difference continued to be directly relevant to performance in adulthood and tried to explain why that should be so, even though sportsmen were no longer competing in year-groups.
- (c) A good many candidates recognized that the blogger's comments were an ineffective response to the research reports, either because she was generalising from a single pair of examples or because the claims made in Sources A and B referred to likelihood or averages rather than to absolutes. However, a lot of candidates wrongly judged that the blogger's comments "disproved" or "refuted" the claims in Sources A and B; in many cases, these judgments were based on a straw man rather than on what Source A actually claimed.
- (d) As indicated on the mark scheme, there were quite a lot of valid points that could have been made by engaging with the sources, and a few candidates achieved good marks on this question, but a large proportion were so out of sympathy with the claim and the research that they rejected them out of hand.

### Question 3

Few candidates showed much sympathy for the ideas in the passage, mostly because they agreed with the counter-argument in paragraph 2, that English has already become the international language.

- (a) A fair number of candidates correctly identified the main conclusion of the passage, but there were also a lot of wrong answers. The most popular wrong answers were, "An international language should be introduced alongside national languages, in order to make communication between different countries easier" (which is part of the introductory scene-setting) and "Of all the languages already devised for international use, Esperanto is the most suitable" (which directly supports the Main Conclusion). A few candidates still summarized the argument instead of analysing it, or offered an analysis of the whole argument instead of identifying the Main Conclusion.
- (b) Many candidates correctly identified at least one of the reasons directly supporting the Main Conclusion, and some identified three. As usual in these questions, all of the correct answers were actually Intermediate Conclusions, i.e. conclusions of mini-arguments. Some candidates drew all their answers to **3(a)** and **3(b)** from a single paragraph, which they should have realized could not be correct.
- (c) Candidates seem gradually to be grasping the nature of this question, although a significant proportion still argued for or against the author's position or evaluated his argument from a literary perspective. Some candidates described strengths in the reasoning, despite the instruction that they should focus on "flaws, unstated assumptions and other weaknesses". Some still mistakenly interpreted the expression "unstated assumptions" as meaning "unsupported statements". Many candidates successfully spotted one or more weak points in the argument, but most of them could not say exactly what was wrong, often presenting a challenge rather than identifying the flaw; some of these achieved good scores by gaining several single marks, awarded for weak attempts at explaining valid points.

The many answers to **3(c)** which were not credited included: the author should have supported his argument by providing statistics; he contradicts his conclusion by admitting that English has already become the international language; he deals well with counter-arguments; he should have considered counter-arguments; the argument is one-sided; the author is biased in favour of

Esperanto; he fails to give any reasons or evidence in support of his claims; the argument is based more on opinion than facts; the author does not give enough information about Esperanto.

- (d) By far the majority of candidates chose to support the claim, but those who opposed it tended to present more persuasive arguments. Some of the reasons frequently deployed in favour of the claim strictly implied that some, rather than all, children should study a foreign language at School. Other reasons were quite vague. A few candidates made some comments without turning them into an argument, but nearly all provided one or more reasons leading to a conclusion. A high proportion of candidates were awarded 3 marks out of 5 for this question, but significant numbers produced structured arguments, which received 4 or 5 marks. A few candidates misunderstood the nature of the task, and discussed Esperanto instead of the claim which they had been asked to discuss; the claim in **3(d)** always has some relation to the theme of the passage used for **3(a)**, **(b)** and **(c)**, but it is always significantly different.

# THINKING SKILLS

---

Paper 9694/22  
Critical Thinking

## Key messages

- Candidates and teachers should study previous question papers and mark schemes, in order to familiarise themselves with the kinds of answers expected to the various questions.
- Candidates need to keep a clear head when evaluating evidence and arguments and avoid thinking that the fact that they disagree with a conclusion constitutes a valid criticism of the evidence or argument.

## General comments

Candidates seemed to respond particularly well to the issues raised by the questions and were able to tackle them effectively. As in previous papers, 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. However, there seemed fewer such candidates for this paper even though the subject matter may have been in danger of encouraging such an approach. It is still the case that some candidates who do well on the first three parts of **Questions 1** and **2** often spend too little time on **part (d)** where a fuller answer is required. However, again there seemed to be fewer such candidates.

## Comments on specific questions

### Question 1

- (a) This was reasonably well-answered with many candidates seeing that Mrs Tripp's refusal was not necessarily suspicious. Some candidates seemed to think it was Mr Tripp refusing to be interviewed.
- (b) This was answered less well, with many candidates assuming the evidence meant he was a good author. Candidates could only get credit for this point if they limited it to him showing some talent in his youth. The best answers did point out that this was not particularly relevant to the question of his talent as a professional writer. The very best made the point that even if he was talented this was not relevant to the issue of whether he actually *did* write the books.
- (c) The vast majority of candidates were able to gain marks by pointing out the lack of reliability due to bias. Fewer made the more subtle point that she was unlikely to say she had received the letter if she had carelessly lost it. More candidates turned this point around and said that it is unlikely she would have lost it if she had received it, making her denial more reliable. Some thought Constance was a man but this did not matter too much.
- (d) The issues raised seem to engage the candidates, who were fairly evenly split between those who argued Mrs Tripp was the author and those who thought it was a plot hatched up by Donald Tripp's enemies. The source material was reasonably balanced between both of these possibilities. Candidates need to guard against over-speculative answers which go well beyond the information in the sources (e.g. it was a plot by Mrs Tripp who hated her husband). They also need to note the reference to "plausible alternative scenarios" which means they should consider the other point of view and give reasons for their rejecting it. The best answers considered the evidence for and against in clearly separated sections before coming to a decision. This was a more effective approach than dealing with each of the sources in turn.

## Question 2

- (a) This discriminated well with good candidates seeing that the inhabitants of New York should still be worried. This was intended to be a question about what could be inferred but some candidates made points by referring to evidence in the other sources. They were able to get some credit for this.
- (b) Candidates were split about 50/50 on this, with those who recognised it as an argument being just a majority. Such candidates usually gained the second mark by accurately referring to the structure of the argument, particularly the conclusion.
- (c) Most candidates correctly interpreted this question as referring to the difference between Kolkhata and Bangkok on the one hand and New York and Tokyo on the other. A minority of candidates unfortunately thought the question referred to the overall increase in population at risk.
- (d) A pleasing number of candidates recognised this was a rather overdrawn conclusion and argued that the sources suggested a rather more modified conclusion, e.g. some parts of cities would become uninhabitable. Few took issue with the rather over-optimistic views expressed in Source D which would not necessarily rule out cities becoming uninhabitable. Rather more candidates were making out a case referring to the sources as necessary rather than simply going through the sources in turn. This meant more Level 2 and 3 answers.

## Question 3

- (a) Only a minority of candidates correctly identified the conclusion. Rather more identified one of the 1 mark answers, usually "Horse racing is not cruel".
- (b) This was very poorly answered with very few candidates identifying even one of the reasons. The reasons to be identified here are those that directly support the main conclusion.
- (c) Candidates who understood the nature of the exercise did well and managed to compensate for poor performance on **Question 1(b)**. There were many 4 and 5 mark answers. However, there are still a significant number of candidates who do not understand the requirement in this question to evaluate the reasoning, usually by spotting flaws and unstated assumptions. They tended to comment on the style of the argument or challenged the propositions. This often meant no marks could be credited.
- (d) Candidates seemed engaged with this conclusion and put forward interesting and often detailed reasons to support their conclusion. Most candidates seemed to be against zoos. A minority of candidates seemed to think the converse of the proposition was that *all* animals should be kept in zoos. Others seemed to think opposing the proposition would mean dangerous wild animals roaming the streets.

# THINKING SKILLS

---

Paper 9694/23  
Critical Thinking

## Key messages

- Candidates and teachers should study previous question papers and mark schemes, in order to familiarise themselves with the kinds of answers expected to the various questions.
- Candidates need to keep a clear head when evaluating evidence and arguments and avoid thinking that the fact that they disagree with a conclusion constitutes a valid criticism of the evidence or argument.

## General comments

Most candidates seemed to respond particularly well to the issues raised by the questions and were able to tackle them effectively. As in previous papers, 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. However, there seemed fewer such candidates for this paper. It is still the case that some candidates who do well on the first three parts of **Questions 1** and **2** often spend too little time on **part (d)** where a fuller answer is required. However, again there seemed to be fewer such candidates.

## Comments on specific questions

### Question 1

- (a) This was not answered well. Candidates needed to assess the significance of this information, in the light of the two possible explanations for the guests being ill that were suggested by the sources. These were food hygiene lapses v. overindulgence. The fact that this person was ill in spite of not overindulging suggests food hygiene was the issue. Very few candidates got this point. Marks were mainly gained for making simple points about reliability along the lines of ability to see and neutrality. Answers that were confined to reliability were capped on 1 mark.
- (b) Whilst being answered better, surprisingly few candidates saw the full degree of significance of this information in contradicting the statements of Mr Singh and the Event Manager *and* offering a convincing explanation of why the guests were ill. Some candidates talked about reliability which was not relevant here. Better candidates saw the significance of a shortage of staff due to illness in terms of possible connections with the guests being ill through cross infection.
- (c) Again, this was not done particularly well, with many candidates only achieving 1 mark for pointing out the confirmation the source gave to the unwise mixing of foods theory of the Events Manager. A minority of better answers saw the lack of *direct* relevance of queue lengths to questions of hygiene, unless it *indirectly* suggested overall incompetence which might extend to food hygiene. Several candidates thought it was a member of staff eating the prawns and ice cream. Others thought staff were responsible for serving this mixture.
- (d) The issues raised seem to engage the candidates and many answers managed to compensate somewhat for rather poor performance in parts (a), (b) and (c). There was some confusion over what constituted 'poor food hygiene', with some candidates seeing the possibility of a power failure exonerating Mr Singh and his staff. However, even these candidates did see the significance of the information in Source F, though only a minority of candidates overall pointed out that we do not know if the power cut was at the time of the event or even whether the event was affected. As usual, over-speculative answers occurred with one candidate claiming Mr Singh owned the Bahan Power Company. Others saw some sort of a plot to sabotage Mr Singh's business. Candidates must make sure that they base their answers on the information they have in the sources.

## Question 2

- (a) This question was answered well with 'requirement of the law' being a favourite response, followed by 'other uses for a lifeboat', especially related to other reasons for needing to abandon the ship.
- (b) Only a minority of answers managed to move beyond the obvious parallels to point out that modern technology would mean a more rapid response and that this could well lead to more people being rescued than in the case of the Titanic.
- (c) There were many good answers to this question though some candidates suggested making the positioning of the lifeboats lower would make it harder to board them without explaining why. Such answers could not be credited
- (d) Many candidates managed to achieve Level 2 answers here and, along with reasonable answers to the other parts, overall marks for **Question 2** were somewhat higher than in previous series. This helped compensate for slightly poorer performance in **Question 1**. The best answers managed to explore the issues in separating out the problems of loading and launching lifeboats from more general problems of an emergency at sea, such as remoteness of location and corridors becoming shafts.

## Question 3

- (a) A majority of candidates managed to get at least 1 mark here with a pleasing number achieving the 2 mark answer.
- (b) Whilst identifying 3 reasons was rare, many candidates identified 2 reasons. Paragraph 4 proved particularly problematic, with many candidates choosing the first sentence as the reason.
- (c) Many candidates saw the problem of generalisation from the Edmonton Mall and questioned why entertainment was sinister or why malls could not simply be *adding* functions rather than *replacing* shops. This meant there were a reasonable number of 3–4 mark answers. Most candidates found it difficult to move beyond these points so 5 mark answers were somewhat rarer. There was again a pleasing increase in the number of candidates who were engaged in the correct exercise of evaluation here, though there is still a significant number of candidates who do not tackle the reasoning in the passage.
- (d) Candidates engaged very well with this conclusion and were equally split between those who argued for and those who argued against on-line shopping. In both cases, they managed to put forward interesting and often detailed reasons to support their conclusion. Arguments often consisted of several independent reasons, which meant maximum marks were rare, but such arguments achieved 4 marks.

# THINKING SKILLS

---

Paper 9694/31

Problem Analysis and Solution

## Key message

Candidates could score more highly by improving the quality of their explanations: the systematic listing of options, the use of trial and improvement methods and the use of words to support numerical calculations are all key aspects of delivering an extended answer to a problem.

## General comments

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. Mathematical preparation for these questions needed to go no further than knowledge of multiples, square numbers (both for **Question 1**) and fractions (**Question 2**).

Many questions required only a brief, unsupported answer; others required an explanation or a demonstration. These latter questions did not all require verbal explanations (a demonstration can simply involve a carefully selected series of calculations), but some did. A number of candidates struggled to find the balance between precision (a numerical answer is normally required if it is possible) and clear explanation (some words are normally needed to make it clear what the significance of a calculation is).

There was a great incentive for candidates to show their working, even if they were unable to furnish a complete solution: there were 25 marks available for clear and sensible working which led to an incorrect final answer.

Although some candidates clearly suffered because they misallocated their time, most attempted all four questions. Candidates' answers to **Question 3** tended to garner the most marks, and **Question 2** delivered the least, in general.

## Comments on specific questions

### **Question 1**

This question required candidates to experiment with lists of consecutive multiples, and their squares, and fit certain requirements regarding their sums. It proved fairly easy to generate erroneous lists of viable-looking widths (such as 4, 6, 8, 10, ...), and this hampered some candidates. The lists themselves were awkward to organise (growing according to their starting number and their length, summing in a non-linear fashion), and this was the key to most of the questions. This question required complete lists which fitted certain conditions, for 7 of the 10 marks, and any omissions or unwanted inclusions prevented marks from being gained.

A few candidates incorrectly took the statement, "the width of the innermost doll is always a whole number of centimetres", (with its plural-looking unit) to imply that the innermost doll could not be 1 cm.

- (a) This question prompted candidates to consider all the valid lists which included 20 cm. The question actually asked for the possible number of dolls – but those who simply offered a list of lists were not penalised. Many candidates realised that this was simply equivalent to 'list the factors of 20'.
- (b) Finding the weight of the heaviest doll with an outermost width of 6 cm required a tactical decision (which list would give the heaviest doll, terminating at 6 cm), and a careful calculation (the summing



of the resulting square numbers). Many candidates failed to do this correctly – mostly because of a tactical error (often yielding the answer 56 g, even though an innermost doll of width of 1 cm was used in **(c)** and **(d)**).

- (c)** In listing the viable pairs of dolls, it was fairly easy to be systematic here – 1 and 2, 2 and 4, 3 and 6, 4 and 8. These had to be converted into weights – which taxed a few candidates (doubling some numbers rather than squaring), but on the whole this was done well. Candidates were not told how many such pairs were needed, and clearly this was part of the decision process – although those who overshot were not penalised here. A substantial number of candidates offered one or two possible pairs, perhaps not realising that a complete list was needed.
- (d)** Tackling this question systematically required candidates to be alert to both aspects of how the lists grew (starting number and length). It proved to be easy to omit at least one of the lists, especially when this process was combined with the squaring of the weights.
- (e)** As with **(d)**, a systematic approach to this was awkward – but candidates did not have to go far beyond the weights they had found to find some that fitted the criterion. Many candidates appeared to have trouble with the idea that the dolls “might not be distinguished by their weights”. The question is a natural conclusion to the investigation (trying to work out how many dolls are in each set, by weighing) – candidates should use the clues from the rest of the question where possible, if they are unsure exactly what is being asked.

## Question 2

This question invited candidates to investigate the interlocking regulations which govern voting in France – which were worthy of careful scrutiny, because of their dependence on both the proportion of voters and the proportion of those eligible to vote. This distinction was carefully defined in the text, but some candidates were clearly confused. It would certainly be recommended that candidates explicitly identify key terms (such as electorate), and attempt to elucidate their meanings and implications before launching into the question. 6 of the 10 marks were awarded for drawing inferences from the data (What does this indicate? What difference would this have made?) and candidates should be prepared to quantify these inferences (i.e. give a precise statement of the numerical implications) where possible.

- (a)** This question required candidates to offer three numbers which fitted the requirements – and was dealt with fairly well. Most candidates opted for variations on the theme of equal proportions (such as 33, 33, 34) or the extreme case (25, 25, 50). A number of candidates simply offered the answer “ $\frac{1}{3}$ ”, which was not in the format requested by the question (“percentages”). Luckily they were credited on this occasion, but such generosity is not guaranteed, and candidates are advised to carefully check what format the answer should be in.
- (b)(i) and (ii)** Many candidates appeared to confuse the electorate and the candidates at this point, giving answers derived from 30 000 rather than the ‘eighth of the electorate’ (or tenth) restriction. The boundary values were to be respected in this question, and hence both marks were only awarded for 7 and 9. Those who failed to appreciate this – the majority of candidates who tackled it with an appropriate method – were only awarded 1 mark.
- (c)(i)** A clear reference to the ‘at least a quarter of the electorate’ rule was all that was needed here, without additional material. This was delivered by most candidates.
  - (ii)** Most candidates did not offer quantitative answers here, and hence scored 0 marks. This does seem to reveal a mistaken expectation about what was needed, since little more than a precise numerical statement of the rule was needed to gain 1 mark: 10 502 is not at least a quarter of the electorate, or the electorate is more than four times 10 502 (42 008). Very few candidates managed this successfully.
  - (iii)** This question was dealt with more competently than **(c)(ii)**.
- (d)** It was expected that candidates would calculate the new numbers of votes, and then re-check how the rules applied. A small proportion of candidates managed this. Some evidence of the key comparison (finding a quarter of the electorate) was deemed necessary for candidates to gain both marks.

### Question 3

This question required candidates to investigate the relationship between certain flows between interconnected compartments, and the quantities in those compartments. The rotation of the compartments effectively introduced two types of flow, and candidates needed to track the flows carefully in order to experiment with the variety of situations posed. The numbers involved were all integers, and less than 20, hence unlikely to prompt arithmetic errors.

As mentioned above, this question delivered the most marks, on average, and most candidates submitted the right kind of answers to all parts.

- (a) This was answered well by most candidates. No restrictions were placed on the format of the answer, and any reasonable format was credited (ordered or unordered).
- (b) (i) and (ii) Candidates were expected to complete a full rotation of the compartments, following the flows shown in the diagram, and about a half of the candidates managed to do this without error. The most common mistakes appeared to involve either failing to rotate the tray at all, or moving the balloons too frequently. Although a diagram is given in this question to facilitate the investigation, candidates need to decide how to track the changes, and most of those who slipped up here appeared to be attempting to complete the turns and flows in their head.
- (c) Candidates were expected to rotate the tray offered in the left-hand diagram, and then manipulate the flows to achieve the right-hand diagram. This was accomplished successfully by a healthy proportion of candidates. Partial marks were available for those who managed to arrange flows which led to certain sectors of the tray being correct.
- (d) This task required candidates to experiment with quantities which would lead to a stable situation – a reverse of the situation in (c). It was less clear how to approach this systematically – and those candidates who did tackle this question appeared to approach an answer by trial and improvement. This question caused significantly more difficulty than (c).
- (e) This question involved a similar task to (c), and was found more accessible than (d). Candidates were expected to find flows which would lead to a stable situation – and a significant proportion of candidates achieved this.
- (f) This extension to the model involved no new logical elements – just a tray with 6 sectors rather than 4. A number of candidates were clearly put off by the perceived complexity of this; but those who attempted the question tended to get a correct answer.

### Question 4

This question required candidates to master the overlapping scores of a 'made-up' game, and to extract data from the league table that accompanied it. The scoring system required careful analysis of complex data, resulting from the fact that different strokes earned different numbers of points, depending on who was the deliverer and who the recipient. The alternating deliverer, and the repeated values in the scoring table, should have prompted candidates to take particular care, and exercise some thought in how they recorded the scores. Many did not manage to do this – and the question attracted proportionately less marks than any of the others.

- (a) (i) This question required candidates to appreciate the basic terminology of the game (20 strands in a match, alternating deliverers, with maximum scoring points of 5 and 9), and assemble a maximum score. Many candidates failed to appreciate the significance of the alternating deliverer, and offered scores of 180 – or offered a maximum score for a strand (9) rather than a match.
- (ii) This question required a careful treatment of the inequality (having won fewer strands than your opponent), followed by careful consideration of the possible scores. Very few candidates attempted to list them in order to check that they were offering viable scores. Only a small number of candidates managed to score 2 marks here.
- (b) This question moved away from the intricacies of the scoring mechanism, to the Order of Play, and was tackled far more competently by the majority of candidates. The question could be approached by a straightforward elimination of those who appeared in all three lists.

- (c) This question required candidates to scrutinise the summary of results, in search of the matching scores – a task that was less easy to ‘track’ on the question paper, and which fewer candidates managed without error. It is hoped that candidates will have had a chance to experience such tables, and to manipulate them, prior to the exam.
- (d) The table which candidates had to analyse in answering this question required some careful unpicking, in order to deduce how many times the two players had each acted as deliverer and recipient. The 20 strands are recorded in the table, and only the 5 point score is ambiguous – but care is needed in tallying up who must have been delivering as each point was won. A full explanation was asked for here, and very few candidates managed this. Candidates should be sensitised to this ‘command word’ in questions – it alerts them to the fact that marks will be put aside for clarity of explanation, as well as the correct judgment. In this question, no marks were available for those who offered a judgment without any comment.
- (e) This question invited candidates to propose a series of 6 strands which would lead to a score of 33–0. These had to include 3 as deliverer and 3 as recipient. Many candidates were able to offer viable suggestions (acquiring partial marks), but few identified the correct solution. This type of question can be difficult to attack systematically; but candidates are encouraged to set out their working clearly nonetheless. In this case, a recorded consideration of possible (winning) scores would have allowed candidates to see how to improve upon their trials, and hence have a better chance of identifying the correct solution.
- (f) Candidates needed to calculate the scores of the most successful players apart from Brow. This was a 3 mark question, and so candidates should expect it to carry partial marks for (appropriate) working, and to be fairly demanding. Disappointingly few candidates left explicit calculations to credit, and most just offered a couple of names (which were normally incorrect). It is likely that these candidates were struggling to complete the paper within the time limit, in which case such behaviour is understandable. Given time, it is hoped that candidates will choose a sensible strategy and attempt to show their progress in the answer booklets.

# THINKING SKILLS

---

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

## Key message

Candidates could score more highly by improving the quality of their explanations: the systematic listing of options, the use of trial and improvement methods and the use of words to support numerical calculations are all key aspects of delivering an extended answer to a problem.

## General comments

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. Mathematical preparation for these questions would have benefited from some exposure to spatial reasoning (both in terms of compass directions in **Question 1**, and the 3D geometry involved in **Question 3**), and some appreciation of probabilities (for **Question 3**).

Many questions required only a brief, unsupported answer; others required an explanation or a demonstration. These latter questions did not all require verbal explanations (a demonstration can simply involve a carefully selected series of calculations), but many did. And a number of candidates struggled to find the balance between precision (a numerical answer is normally required if it is possible) and clear explanation (some words are normally needed to make it clear what the significance of a calculation is).

There was a great incentive for candidates to show their working, even if they were unable to furnish a complete solution: there were 22 marks available for sensible working which led to an incorrect final answer.

Although some candidates clearly suffered because they misallocated their time, most attempted all four questions. Candidates' answers to **Question 4** tended to garner the most marks, and **Question 1** delivered the least, in general.

## Comments on specific questions

### **Question 1**

This question involved modelling the behaviour of a ball inside a simple maze, as the maze is altered according to certain simple principles. The question could be divided into two separate activities – an initial analysis of what mazes were possible, given the restrictions on where the bungs could be placed, followed by experimental consideration of how the ball would be affected by tilting in one of the 8 compass directions. The first six marks were won by orderly consideration of the options available for the 'bungs', and many candidates did list the options they considered. The majority appeared to have completed these tasks mentally, with mixed results.

- (a) Many candidates misjudged this question, and showed no working – and were awarded 0 marks. The careful description of what is meant by "blocking off any sections of the maze", given in the paragraph above ("no white space unreachable") needed to be applied to a 'blank' maze. The first requirement of many such problem-solving questions is to ensure that any problem-specific terms are being used correctly, and this is what this question is testing.
- (b) This built on the understanding developed in (a), but was dealt with better than (a) because candidates were invited to offer a list of options – and hence a number of candidates were able to gain partial marks. The question allowed candidates to appreciate which combinations of bungs were permissible for the next stage of the question.

- (c) Many candidates appreciated that the nine pockets which exist in the ‘blank’ maze, could be reduced to six if bungs were placed in three of them. This yielded the minimum numbers of pockets. The task of finding the maximum was more deceptive – the tempting solution, which was to place three bungs together in a corner and hence create two new pockets was not maximal; two bungs in D1 and D2 (or D4 and D5) replaced one pocket with three new ones, and could be combined with another bung elsewhere to create 10 in total. Very few candidates identified this.
- (d) Most candidates offered a possible solution to this, and gained at least 1 mark. 1 mark was awarded for a viable starting point requiring at least two tilts, as long as it was accompanied by directions which would lead to the ball’s escape. There were a number of such positions, but only one which required three tilts. The diagram below shows the minimum number of tilts needed from each starting square:

	3		2		1	
2	2	2	1	1	1	1
			1		1	
	1	1	1	1	1	2
	1				2	

- (d) Although tackled less successfully than (c), many candidates did manage to offer positions for two bungs and the ball which required at least three tilts – and hence gained partial marks here.

## Question 2

This question required candidates to appreciate the relationship between the table (showing states, inputs and ensuing states) and the door codes which might unlock a door. The questions involved an understanding of how the table encoded the information, and of how certain aspects of the table might produce certain types of ‘behaviour’ in the codes. The majority of candidates appreciated the basic structure of the table, but few were able to apply this creatively and without error in part (d).

- (a) This question offered candidates a chance to check that they understood how the table worked, offering the final outcome of Mark entering the code CBCA. Most candidates were able to offer the two intervening states successfully, and gain 2 marks. Those who did not accomplish this often appeared to be attempting to fit the question to their expectations of what the door might be doing (for instance, moving progressively through the states).
- (b) Almost all those who completed (a) successfully also managed (b), which represented the first step of applying the rule verified in (a).
- (c) This question required candidates to consider what the implication of a column of 1s was: most candidates appreciated that the explanation must link the 1s to the intended door code, but a few candidates offered circular answers such as “*It is so that the door always returns to State 1 after B is pressed*”, and were not credited. Any explanation which referred to B not being in the code, or B possibly acting as a ‘reset’ button, were credited.
- (c) This question tested both understanding and the articulation of that understanding – quite a few candidates were able to see (or guess) that the peculiar behaviour of ‘A’ must be due to the repeated ‘C’, but did not articulate this clearly. The best attempts at this question used very specific examples – such as “*Pressing A while in State 4 must have been achieved by entering the code CACA – in which case the first two letters have been entered correctly...*”
- (e) Almost all those candidates who attempted this question gained at least 1 mark, by filling in the mechanism’s response to the appropriate ‘correct pathway’. A small number also managed to fill in the background 1s, and some of the first column. Very few completed the whole table correctly.

### Question 3

This question required candidates to consider how the different aspects of a regular Platonic solid could be used to differentiate between a number of players fairly. Success at this question did not benefit from any profound knowledge of probability, or Platonic solids, but a careful consideration of how edges, corners and orientations could be inventively used. The question did also test candidates' ability to articulate their methods for using the dice with appropriate precision – something that some candidates clearly struggled with.

- (a) The majority of candidates were able to describe the method for choosing between three people with a cubical die. A significant minority did fall into the trap offering each person one number and discarding the rest – which although fair does not achieve its aim reliably in one throw of the die.
- (b) This question requires a creative use of the cube, following on from the examples given by Alexander and Zoe. Not many candidates managed to offer an unambiguous description of their method, and so few were awarded 2 marks as a result. Many candidates offered answers that were woefully brief, or which failed to clearly to identify how the edges were to be allocated.
- (c) (i) This question required candidates to methodically list the options achievable, subject to the restrictions available on adjacent sides of a die. Many candidates did not list the options on paper, made an error, and scored 0 marks. There were generous marks available for those who attempted to list them. The correct list of number pairs could be achieved by laying out a complete list of (30) pairs, and then deleting those which involved numbers adding up to 7.
  - (ii) This question favoured those who had already made systematic lists of the appropriate pairings – and was tackled with greater than competence than part (i).
- (e) Most candidates who attempted this question answered it successfully.
- (e) This was a 3 mark question, and a precise, clearly-articulated answer was expected. A number of candidates appreciated that the pentagonal faces of the dodecahedron must be the key to the identification of one player from five, but could not articulate how to use it. This tended to gain 1 mark. Those who explicitly identified the five corners of the face on top, or the 30 edges of the shape in total, or the 20 vertices in total gained at least 2 marks. And a clear statement of how one of these aspects was to be identified was needed for the third mark – a simple way of doing this was offered by Alexander at the beginning.
- (f) (i) and (ii) Candidates tended to achieve both or neither of these marks – and many candidates appeared to be exhausted by the expectations of the question by this point. Surprisingly few stated the number of edges of the dodecahedron, which was the easiest way to an answer here (combined with Zoe's method, given in the example).
- (g) The need to apply the independent probabilities achievable from each die, and therefore to factorise 45, was not obvious to many candidates – a few candidates worked back from 45 to speculate on how the two dice might be used, and gained a mark regardless of whether they were able to present a workable plan.

### Question 4

This question involved the investigation of a simple competitive game, and was dependent on the careful checking of what remaining pieces were available. The question demanded explanations of what was possible in the game, and hence an appreciation of when something was proved, regardless of what the opponent might do. This subtlety to the explanations would be worth drawing to the attention of candidates – the distinction between what is possible in a game, and what is inevitable, underlies a number of such problem-solving scenarios, and is a useful distinction to master.

- (a) (i) and (ii) Most candidates managed to calculate the correct scores for Xavier and Yvette; a few offered a score of 11 for part (ii), presumably by discounting the middle column, as a result of mistakenly aligning the pair of 3s or 4s with Yvette.
- (b) This question was answered well, with almost all candidates choosing to present a pair of 4s, 5s and 6s in order to maximise Yvette's score, and appropriate pairs on the rows to minimise Xavier's score. The most common errors were to include a third number in the grid accidentally, or to initiate

a grid which was not possible to complete without a 1 (such as is produced by placing 456 in each of the first two columns).

- (c) The explanation required by this question was probably the most demanding of all the verbal questions on the paper – some attempt to address the inevitability of a non-zero score was needed to be awarded 2 marks. Most were awarded a mark of 1 for offering evidence of the possibility of a non-zero score: by drawing an example grid, or appealing to what happens when two columns are lined up with ‘abc’ in each. Those who could not even manage this – often appealing vaguely to the fact that there were an odd number of squares in the grid – did not collect any marks. The best explanations considered more than one arrangement of numbers in the grid, or attempted to supplement their examples with consideration of alternatives.
- (d) As with (c), this question required a sense of the inevitability of the 3 and two 4s being left in the bag. Two marks were awarded for showing it was possible (a grid which yielded a score of 10–10, and used all the other numbers). However, only a few appreciated that the question was asking for a demonstration that this was the only possible outcome (which was needed to be awarded 3 marks).
- (e) A single drawing of the grid was all that was needed to complete this question correctly – and most candidates found it easier than (d). It was difficult to analyse how candidates tackled the question (by trial and improvement, or by logically working back from the final score, and the completion of the middle column). The most common error was to accidentally include a third occurrence of a number in the grid.
- (f) This question explicitly asked for a demonstration that Yvette could be sure of her victory – but most solutions only showed that it was possible (not certain). Some explicit attempt to analyse the best moves that Xavier could play was necessary to accomplish this. A significant number of candidates were tempted to offer qualitative arguments to the question (“*If she places the 6 at the top, she will have highest scoring numbers on the grid, and Xavier will have lost a row, and so she will definitely score far more than Xavier...*”) – in response to this, the general rule should be that any argument required in Paper 3 should be attempted using precise numerical examples if at all possible.

# THINKING SKILLS

---

Paper 9694/33

Problem Analysis and Solution

## Key message

Candidates could score more highly by improving the quality of their explanations: the systematic listing of options, the use of trial and improvement methods and the use of words to support numerical calculations are all key aspects of delivering an extended answer to a problem.

## General comments

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. Mathematical preparation for these questions needed to go no further than knowledge of multiples, square numbers (both for **Question 1**) and fractions (**Question 2**).

Many questions required only a brief, unsupported answer; others required an explanation or a demonstration. These latter questions did not all require verbal explanations (a demonstration can simply involve a carefully selected series of calculations), but some did. A number of candidates struggled to find the balance between precision (a numerical answer is normally required if it is possible) and clear explanation (some words are normally needed to make it clear what the significance of a calculation is).

There was a great incentive for candidates to show their working, even if they were unable to furnish a complete solution: there were 25 marks available for clear and sensible working which led to an incorrect final answer.

Although some candidates clearly suffered because they misallocated their time, most attempted all four questions. Candidates' answers to **Question 3** tended to garner the most marks, and **Question 2** delivered the least, in general.

## Comments on specific questions

### **Question 1**

This question required candidates to experiment with lists of consecutive multiples, and their squares, and fit certain requirements regarding their sums. It proved fairly easy to generate erroneous lists of viable-looking widths (such as 4, 6, 8, 10, ...), and this hampered some candidates. The lists themselves were awkward to organise (growing according to their starting number and their length, summing in a non-linear fashion), and this was the key to most of the questions. This question required complete lists which fitted certain conditions, for 7 of the 10 marks, and any omissions or unwanted inclusions prevented marks from being gained.

A few candidates incorrectly took the statement, "the width of the innermost doll is always a whole number of centimetres", (with its plural-looking unit) to imply that the innermost doll could not be 1 cm.

- (a) This question prompted candidates to consider all the valid lists which included 20 cm. The question actually asked for the possible number of dolls – but those who simply offered a list of lists were not penalised. Many candidates realised that this was simply equivalent to 'list the factors of 20'.
- (b) Finding the weight of the heaviest doll with an outermost width of 6 cm required a tactical decision (which list would give the heaviest doll, terminating at 6 cm), and a careful calculation (the summing



of the resulting square numbers). Many candidates failed to do this correctly – mostly because of a tactical error (often yielding the answer 56 g, even though an innermost doll of width of 1 cm was used in **(c)** and **(d)**).

- (c)** In listing the viable pairs of dolls, it was fairly easy to be systematic here – 1 and 2, 2 and 4, 3 and 6, 4 and 8. These had to be converted into weights – which taxed a few candidates (doubling some numbers rather than squaring), but on the whole this was done well. Candidates were not told how many such pairs were needed, and clearly this was part of the decision process – although those who overshot were not penalised here. A substantial number of candidates offered one or two possible pairs, perhaps not realising that a complete list was needed.
- (d)** Tackling this question systematically required candidates to be alert to both aspects of how the lists grew (starting number and length). It proved to be easy to omit at least one of the lists, especially when this process was combined with the squaring of the weights.
- (e)** As with **(d)**, a systematic approach to this was awkward – but candidates did not have to go far beyond the weights they had found to find some that fitted the criterion. Many candidates appeared to have trouble with the idea that the dolls “might not be distinguished by their weights”. The question is a natural conclusion to the investigation (trying to work out how many dolls are in each set, by weighing) – candidates should use the clues from the rest of the question where possible, if they are unsure exactly what is being asked.

## Question 2

This question invited candidates to investigate the interlocking regulations which govern voting in France – which were worthy of careful scrutiny, because of their dependence on both the proportion of voters and the proportion of those eligible to vote. This distinction was carefully defined in the text, but some candidates were clearly confused. It would certainly be recommended that candidates explicitly identify key terms (such as electorate), and attempt to elucidate their meanings and implications before launching into the question. 6 of the 10 marks were awarded for drawing inferences from the data (What does this indicate? What difference would this have made?) and candidates should be prepared to quantify these inferences (i.e. give a precise statement of the numerical implications) where possible.

- (a)** This question required candidates to offer three numbers which fitted the requirements – and was dealt with fairly well. Most candidates opted for variations on the theme of equal proportions (such as 33, 33, 34) or the extreme case (25, 25, 50). A number of candidates simply offered the answer “ $\frac{1}{3}$ ”, which was not in the format requested by the question (“percentages”). Luckily they were credited on this occasion, but such generosity is not guaranteed, and candidates are advised to carefully check what format the answer should be in.
- (b)(i) and (ii)** Many candidates appeared to confuse the electorate and the candidates at this point, giving answers derived from 30 000 rather than the ‘eighth of the electorate’ (or tenth) restriction. The boundary values were to be respected in this question, and hence both marks were only awarded for 7 and 9. Those who failed to appreciate this – the majority of candidates who tackled it with an appropriate method – were only awarded 1 mark.
- (c)(i)** A clear reference to the ‘at least a quarter of the electorate’ rule was all that was needed here, without additional material. This was delivered by most candidates.
- (ii)** Most candidates did not offer quantitative answers here, and hence scored 0 marks. This does seem to reveal a mistaken expectation about what was needed, since little more than a precise numerical statement of the rule was needed to gain 1 mark: 10 502 is not at least a quarter of the electorate, or the electorate is more than four times 10 502 (42 008). Very few candidates managed this successfully.
- (iii)** This question was dealt with more competently than **(c)(ii)**.
- (d)** It was expected that candidates would calculate the new numbers of votes, and then re-check how the rules applied. A small proportion of candidates managed this. Some evidence of the key comparison (finding a quarter of the electorate) was deemed necessary for candidates to gain both marks.

### Question 3

This question required candidates to investigate the relationship between certain flows between interconnected compartments, and the quantities in those compartments. The rotation of the compartments effectively introduced two types of flow, and candidates needed to track the flows carefully in order to experiment with the variety of situations posed. The numbers involved were all integers, and less than 20, hence unlikely to prompt arithmetic errors.

As mentioned above, this question delivered the most marks, on average, and most candidates submitted the right kind of answers to all parts.

- (a) This was answered well by most candidates. No restrictions were placed on the format of the answer, and any reasonable format was credited (ordered or unordered).
- (b) (i) and (ii) Candidates were expected to complete a full rotation of the compartments, following the flows shown in the diagram, and about a half of the candidates managed to do this without error. The most common mistakes appeared to involve either failing to rotate the tray at all, or moving the balloons too frequently. Although a diagram is given in this question to facilitate the investigation, candidates need to decide how to track the changes, and most of those who slipped up here appeared to be attempting to complete the turns and flows in their head.
- (c) Candidates were expected to rotate the tray offered in the left-hand diagram, and then manipulate the flows to achieve the right-hand diagram. This was accomplished successfully by a healthy proportion of candidates. Partial marks were available for those who managed to arrange flows which led to certain sectors of the tray being correct.
- (d) This task required candidates to experiment with quantities which would lead to a stable situation – a reverse of the situation in (c). It was less clear how to approach this systematically – and those candidates who did tackle this question appeared to approach an answer by trial and improvement. This question caused significantly more difficulty than (c).
- (e) This question involved a similar task to (c), and was found more accessible than (d). Candidates were expected to find flows which would lead to a stable situation – and a significant proportion of candidates achieved this.
- (f) This extension to the model involved no new logical elements – just a tray with 6 sectors rather than 4. A number of candidates were clearly put off by the perceived complexity of this; but those who attempted the question tended to get a correct answer.

### Question 4

This question required candidates to master the overlapping scores of a ‘made-up’ game, and to extract data from the league table that accompanied it. The scoring system required careful analysis of complex data, resulting from the fact that different strokes earned different numbers of points, depending on who was the deliverer and who the recipient. The alternating deliverer, and the repeated values in the scoring table, should have prompted candidates to take particular care, and exercise some thought in how they recorded the scores. Many did not manage to do this – and the question attracted proportionately less marks than any of the others.

- (a) (i) This question required candidates to appreciate the basic terminology of the game (20 strands in a match, alternating deliverers, with maximum scoring points of 5 and 9), and assemble a maximum score. Many candidates failed to appreciate the significance of the alternating deliverer, and offered scores of 180 – or offered a maximum score for a strand (9) rather than a match.
- (ii) This question required a careful treatment of the inequality (having won fewer strands than your opponent), followed by careful consideration of the possible scores. Very few candidates attempted to list them in order to check that they were offering viable scores. Only a small number of candidates managed to score 2 marks here.
- (b) This question moved away from the intricacies of the scoring mechanism, to the Order of Play, and was tackled far more competently by the majority of candidates. The question could be approached by a straightforward elimination of those who appeared in all three lists.

- (c) This question required candidates to scrutinise the summary of results, in search of the matching scores – a task that was less easy to ‘track’ on the question paper, and which fewer candidates managed without error. It is hoped that candidates will have had a chance to experience such tables, and to manipulate them, prior to the exam.
- (d) The table which candidates had to analyse in answering this question required some careful unpicking, in order to deduce how many times the two players had each acted as deliverer and recipient. The 20 strands are recorded in the table, and only the 5 point score is ambiguous – but care is needed in tallying up who must have been delivering as each point was won. A full explanation was asked for here, and very few candidates managed this. Candidates should be sensitised to this ‘command word’ in questions – it alerts them to the fact that marks will be put aside for clarity of explanation, as well as the correct judgment. In this question, no marks were available for those who offered a judgment without any comment.
- (e) This question invited candidates to propose a series of 6 strands which would lead to a score of 33–0. These had to include 3 as deliverer and 3 as recipient. Many candidates were able to offer viable suggestions (acquiring partial marks), but few identified the correct solution. This type of question can be difficult to attack systematically; but candidates are encouraged to set out their working clearly nonetheless. In this case, a recorded consideration of possible (winning) scores would have allowed candidates to see how to improve upon their trials, and hence have a better chance of identifying the correct solution.
- (f) Candidates needed to calculate the scores of the most successful players apart from Brow. This was a 3 mark question, and so candidates should expect it to carry partial marks for (appropriate) working, and to be fairly demanding. Disappointingly few candidates left explicit calculations to credit, and most just offered a couple of names (which were normally incorrect). It is likely that these candidates were struggling to complete the paper within the time limit, in which case such behaviour is understandable. Given time, it is hoped that candidates will choose a sensible strategy and attempt to show their progress in the answer booklets.

# THINKING SKILLS

---

Paper 9694/41  
Applied Reasoning

## Key Messages

- The first question in this paper tested the candidates' ability to evaluate claims based on statistical data.
- In **Question 2** candidates had the opportunity to display their ability to analyse the structure of a detailed argument.
- In **Question 3** candidates only gained marks if they made comments about the strength of the reasoning within the document.
- **Question 4** allowed candidates to use a full range of critical reasoning skills in order to construct a reasoned argument using information from the documents.

## General Comments

There were some pleasing signs of improvement in the preparation of some candidates. The different style of **Question 1** seemed to encourage candidates to spend extra time on **Question 4**. Hence, responses to **Question 4** were often longer, and in some cases better, than in previous sessions.

## Comments on Specific Questions

### Question 1

Many candidates found this question more accessible than previous **Question 1s**. Unfortunately, some candidates, despite being instructed to simply "write down true or false", chose to explain their answers in detail. This cost them valuable time which could have been spent thinking about the higher-tariff questions. It was very rare for anyone to score less than two marks.

### Question 2

There were an increased number of candidates actually attempting an analysis, which is a clear indication that some Centres had been preparing candidates well for the examination. On the other hand a minority of candidates still provided a summary or gist, suggesting a lack of preparation. Some candidates seemed unaware that quoting from the text is an appropriate, indeed a required, way to answer this question. The mark scheme this year was less forgiving of candidates who included extra elements in their answers beyond those asked for by the question. The document provided was challenging to analyse and a minority of candidates gained more than half marks; only a handful achieved all six marks. The best and most time-efficient answers copied directly (as far as possible) from the text and presented each element as a bullet point labelled MC, IC, CA, etc.

### Question 3

Again, it was pleasing to see more candidates attempting to evaluate the passage, but a good number continued to see the requirements as listing a series of counter-arguments to points in the passage or even producing a critique of the author's literary style. Those candidates who did attempt to apply their evaluation skills found the question challenging. Well prepared candidates tended to score between two and five marks; candidates who achieved higher marks were in a small minority. Many responses identified the *ad hominem* fallacy in paragraph 4. Other weaknesses were identified by a number of candidates, but they were often badly expressed or incorrectly labelled and were awarded one mark rather than two.

#### Question 4

The majority of candidates secured between 10 and 15 marks on the strength of their having presented an argument towards a conclusion supported by reasons largely lifted from the documents. In order to achieve Level 4, an answer needs to show developed skills in four areas: structure, quality of argument, use of documents and treatment of counter positions. There were few cases in which candidates developed arguments with accomplished structural complexity. As the subject matter, the UN and its Security Council, seemed unfamiliar to many candidates, some found it difficult to use ideas beyond those presented in the documents. With respect to the use of documents themselves, well-prepared candidates did attempt to evaluate the usefulness of sources, but in most cases this did not go much beyond the consideration of bias. In a minority of cases the responses simply described the contribution made by each document to the debate. Such responses tended to gain some credit for use of documents but very little for structure. 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 is likely to get high marks.

# THINKING SKILLS

---

Paper 9694/42  
Applied Reasoning

## Key Messages

- The first question in this paper tested the candidates' ability to evaluate claims based on statistical data.
- In **Question 2** candidates had the opportunity to display their ability to analyse the structure of a detailed argument.
- In **Question 3** candidates only gained marks if they made comments about the strength of the reasoning within the document.
- **Question 4** allowed candidates to use a full range of critical reasoning skills in order to construct a reasoned argument using information from the documents.

## General Comments

There was little evidence of candidates running out of time on this paper. A smaller, but still significant, proportion of candidates are writing answers whose length does not reflect the mark allocation – responses to **Question 1** should be considerably shorter than those to **Question 4**.

The standard of candidates varied and there was evidence that many candidates had been taught some of the language of reasoning and some were familiar with the format of the paper. Indeed, some candidates appeared particularly well-prepared in this regard and answered **Question 4** first, attempting to ensure that the most creditworthy question was not rushed.

## Comments on Specific Questions

### **Question 1**

Surprisingly, this was many candidates' weakest question.

- (a) Candidates were asked to criticise the statistics in the deliberately poor-quality pie chart. However, the majority of candidates attempted to criticise the inferences drawn from the statistics – which is what they were asked to do in part (b). Responses matching all of the points in the mark scheme were seen.
- (b) Candidates did better on part (b), often repeating an answer from part (a) but this time gaining credit for it. Most two-mark responses identified and then explained the lack of representativeness of the sample or the problems with the survey question.

### **Question 2**

Candidates who had been prepared for the examination usually gained some marks for this question, although it was rare to award more than four marks because identification of the main conclusion was more challenging than in some previous papers. As always, some candidates did not understand what was required of the task and attempted to paraphrase, summarize or criticize the argument. Successful candidates identified parts of the text, copied them out and labelled them as MC, IC, etc.

### Question 3

Low- to medium-level evaluation of the argument was reasonably straightforward, and candidates who knew what 'evaluate' meant and who had learned some critical thinking were generally able to score at least three, and often more, marks. Unprepared candidates, who stated a series of counter-assertions to the claims made in the document, were fewer than in previous sessions – around one third of the cohort.

### Question 4

The majority of candidates found the subject matter accessible and most had strong opinions. This helped the weaker candidates construct focused arguments. There was evidence that more candidates were using ideas of their own. While the vast majority of candidates used the documents to support their argument, disappointingly many made no critical comments about the documents they were using, limiting their mark in the 'use of documents' skill area. Conversely, some candidates answered the question by discussing each document in turn rather than structuring their answer as an argument, limiting their mark in the 'structure' skill area. While it was pleasing to see the majority of candidates offer some sort of conclusion, a disappointing number are still writing "I agree" or "I disagree". Candidates should remember that an argument needs a conclusion – and a precise and well-supported conclusion is likely to achieve more marks.

# THINKING SKILLS

---

Paper 9694/43  
Applied Reasoning

## Key Messages

- The first question in this paper tested the candidates' ability to evaluate claims based on statistical data.
- In **Question 2** candidates had the opportunity to display their ability to analyse the structure of a detailed argument.
- In **Question 3** candidates only gained marks if they made comments about the strength of the reasoning within the document.
- **Question 4** allowed candidates to use a full range of critical reasoning skills in order to construct a reasoned argument using information from the documents.

## General Comments

There were some pleasing signs of improvement in the preparation of some candidates. The different style of **Question 1** seemed to encourage candidates to spend extra time on **Question 4**. Hence, responses to **Question 4** were often longer, and in some cases better, than in previous sessions.

## Comments on Specific Questions

### Question 1

Many candidates found this question more accessible than previous **Question 1s**. Unfortunately, some candidates, despite being instructed to simply "write down true or false", chose to explain their answers in detail. This cost them valuable time which could have been spent thinking about the higher-tariff questions. It was very rare for anyone to score less than two marks.

### Question 2

There were an increased number of candidates actually attempting an analysis, which is a clear indication that some Centres had been preparing candidates well for the examination. On the other hand a minority of candidates still provided a summary or gist, suggesting a lack of preparation. Some candidates seemed unaware that quoting from the text is an appropriate, indeed a required, way to answer this question. The mark scheme this year was less forgiving of candidates who included extra elements in their answers beyond those asked for by the question. The document provided was challenging to analyse and a minority of candidates gained more than half marks; only a handful achieved all six marks. The best and most time-efficient answers copied directly (as far as possible) from the text and presented each element as a bullet point labelled MC, IC, CA, etc.

### Question 3

Again, it was pleasing to see more candidates attempting to evaluate the passage, but a good number continued to see the requirements as listing a series of counter-arguments to points in the passage or even producing a critique of the author's literary style. Those candidates who did attempt to apply their evaluation skills found the question challenging. Well prepared candidates tended to score between two and five marks; candidates who achieved higher marks were in a small minority. Many responses identified the *ad hominem* fallacy in paragraph 4. Other weaknesses were identified by a number of candidates, but they were often badly expressed or incorrectly labelled and were awarded one mark rather than two.



#### Question 4

The majority of candidates secured between 10 and 15 marks on the strength of their having presented an argument towards a conclusion supported by reasons largely lifted from the documents. In order to achieve Level 4, an answer needs to show developed skills in four areas: structure, quality of argument, use of documents and treatment of counter positions. There were few cases in which candidates developed arguments with accomplished structural complexity. As the subject matter, the UN and its Security Council, seemed unfamiliar to many candidates, some found it difficult to use ideas beyond those presented in the documents. With respect to the use of documents themselves, well-prepared candidates did attempt to evaluate the usefulness of sources, but in most cases this did not go much beyond the consideration of bias. In a minority of cases the responses simply described the contribution made by each document to the debate. Such responses tended to gain some credit for use of documents but very little for structure. 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 is likely to get high marks.