

THINKING SKILLS

Paper 9694/21 Critical Thinking
--

Key Messages

Teachers and students should use old mark schemes and Principal Examiner's Reports in order to familiarise themselves with the types of answers expected to formats of question which frequently occur. For example, candidates need to know what questions about the "reliability" of evidence mean (as in **Question 1(b)**) and the kind of answer which would be credited.

General comments

Nearly all candidates had time to attempt all questions. Some of them wrote their answers to **Question 3(d)** before attempting **Question 3(c)**, which is a wise strategy, especially if time is short. Others followed a similar strategy by answering **Question 3** before **Question 2**.

Comments on specific questions

Question 1

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

- (a) A full answer to this question would consist of three steps: a valid observation about retired Special Forces Marines; an inference from that observation to Peter Day's behaviour on the occasion in question; and an inference from that behaviour to the reliability of his claim that he did not hit the intruders. Any two of those were sufficient to achieve 2 marks out of 2, and most candidates were awarded full marks. Since the focus of the question was on inferential reasoning, it was possible to achieve 1 or even 2 marks by drawing reasonable inferences from implausible observations. A few candidates made general evaluative remarks about Mr Day's evidence, not related to his former career, but these were not credited.
- (b) There were two valid lines of approach to this question, based on the credibility of the source and the plausibility of the claim, and in order to gain full marks it was necessary to pursue both approaches. Many candidates achieved 1 mark, a fair number 2 marks and a few 3 or 4. The most popular answer, which was not credited, was that anyone who is out and about at 3 am is bound to be up to no good. The most popular correct answer was that Darren Jones should have rung either the police or the doorbell, but relatively few candidates went on to explain that his failure to do so made Darren Jones's explanation for his presence in the house implausible. Not many candidates discussed the vested interest of Darren Jones to exculpate himself from a charge of burglary by lying about his purpose in entering Peter Day's house, although that would have been the obvious approach for anyone who was familiar with questions about the reliability of evidence. Some of those who did recognise the vested interest to appear innocent did not explain that this could cause him to lie.

- (c) Four valid comments could have been made about the relevance of the law textbook to this case, any three of which were sufficient for full marks. Two of those comments referred to what the book did say, and two to what it did not. Many candidates achieved 1 mark, and some 2, but very few saw both sides, which was necessary to identify three points. Almost no one recognised both elements of the negative aspect of the textbook, namely that it neither helped to establish what Mr Day had actually done nor explained what might constitute 'reasonable' force. Some candidates wrongly claimed that Source E would show whether Mr Day used excessive force or not, whereas one of the points on the markscheme was that it did not give this information. A significant minority of candidates summarised what the book said, without relating it to the case of Mr Day and Darren Jones, and were awarded 0 marks.
- (d) Opinions varied amongst candidates as to whether Mr Day did not hit Darren Jones, used reasonable force in self-defence or used excessive force, but the last of these was the most popular. Some candidates who attempted to identify and discuss an alternative either used the same option they had already chosen or identified a hypothesis (such as that Darren Jones was trying to rob Mr Day's house) which was not an alternative to Mr Day's possible guilt. Many candidates thought that for a 68-year-old to hit two young intruders with a weapon and push them downstairs constituted excessive force. Some discussions did not support the conclusion drawn, especially arguing that Mr Day was guilty of assault because he used proportionate force in order to defend himself. Many answers achieved 5 or 6 marks out of 6, by evaluating the sources and/or using inferential reasoning. Some attempted evaluations of the sources lacked justification, *e.g.* by calling one 'implausible' without saying why this was so. Some weak answers claimed either that Mr Day was innocent simply because he denied hitting Darren Jones or that he was guilty of assault just because Darren Jones said he had been hit and thrown down the stairs. Many candidates took Darren Jones's preposterous account of his actions more seriously than it deserved.

Question 2

Candidates appeared to have no difficulty in understanding this issue.

- (a) To perform well on this question, it was necessary to understand what the company's vested interest would tempt them to do, and then to notice that they had not done so, thereby strengthening their reliability. By far the most popular answer was a simple application of vested interest, claiming that the advice on the company's website was probably false, in order to mislead the public into catching colds, and thereby increase sales of their products. A full version of that answer was given 1 mark, because it showed that at least the candidate understood how vested interest **might** work, even though in this case it had not; however, generic answers along these lines, which did not refer to what the passage actually said, were not credited. A mark was also available for pointing out that the reliability was strengthened by the company's expertise. Answers which said that the reliability was not affected, clearly meaning that it was not reduced or compromised, were credited.
- (b) (i) Candidates found it very difficult to explain how the information in Source C about varying degrees of symptoms of colds and flu supported the hypothesis of the researchers in Source B about some people ignoring symptoms. Some of those who took this approach over-stated the connexion, saying that having no symptoms and ignoring symptoms were synonymous; they were not awarded the mark. More candidates, although still not many, gained the second mark available, by pointing out that Source C had no reference to whether people had been chilled or not, which was the focus of Source B.
- (ii) Many acceptable answers to this question were available, and most candidates who focused on the group of participants whose feet had been chilled thought of two such answers, thereby gaining 2 marks. However, many candidates who gave potentially correct answers referred to all the participants in the study who reported cold symptoms, and not specifically to the group whose feet had been chilled, which was the crucial point, and so they were awarded 0 marks.
- (c) A fair number of candidates achieved 1 mark, by identifying either the similarities or the differences between mice and humans, and a few achieved 2 marks, by mentioning both. 1 mark was also awarded to candidates who correctly pointed out that the question referred to becoming infected, whereas the focus of the experiment was the ability to fight against infection. Many answers did not refer to the species being studied, which suggested that candidates may not have noticed that the experiment was conducted on mice, and the findings applied to humans.

- (d) Most candidates disagreed with the claim, and relied largely on Source A to support that judgment. A good number gave a nuanced conclusion, based on the ambiguity of the word “avoid”, arguing that it is possible to reduce, but not to eliminate, the risk of catching a cold. The relevance of the other sources to the claim was not as obvious as in some papers, and some candidates were not credited for their use of sources because they summarised what the sources said without relating them to the claim. Many candidates had no difficulty in claiming both that low temperatures do not make people more likely to catch a cold or flu (Source A) and that they do (Sources B and D), without mentioning the contradiction. Most candidates were unable to relate Source C to the question posed, although some linked it creatively to Source A, to infer that it is impossible to avoid contact with infected people, because many people who are infected with a cold or flu are unaware of it. Some others inferred that as the base of the Common Cold Pyramid shows a substantial proportion of people have no infection despite exposure to the virus, this suggests that is indeed possible to avoid catching a cold.

Question 3

Nearly all candidates seemed engaged by this issue, but some were apparently so antipathetic towards the author’s opinions that they found it hard to evaluate them coolly.

- (a) Most candidates correctly identified the main conclusion, although a significant minority did not. Most of the wrong answers consisted of intermediate conclusions, but some candidates offered the last sentence of paragraph 1, presumably because they had been advised that the main conclusion is often located there. A few candidates offered a précis of the passage instead of identifying the main conclusion.
- (b) Many candidates correctly identified 2 or 3 correct answers to this question. The first sentence of paragraph 5 was the most popular wrong answer. Only a few seemed not to understand the nature of this task.
- (c) A fair number of candidates understood what was expected by this question, and gave at least one correct answer, although as in previous series some still argued against the reasoning instead of evaluating it. Some candidates wasted time attempting to identify and explain strengths in the reasoning, even though the question made it clear that the overall strength should be assessed by identifying “flaws, unstated assumptions and other weaknesses”. Some candidates correctly identified significant unstated assumptions, but many still wrongly interpreted the expression “unstated assumptions” as meaning “unsupported statements”. The most popular correct answers identifying flaws were the slippery slope and restricted options in paragraph 1. Some candidates also spotted the appeal to tradition in paragraph 3 and the inconsistency between paragraphs 1 and 5 concerning the value of a university education. Some candidates wrongly identified the reference to “responsible parents” in paragraph 5 as an *ad hominem* flaw, while others misunderstood the analogy in that sentence and criticised the author for describing the university experience as consisting of “pleasure, luxury and self-indulgence”. As on previous occasions, marks were not awarded for criticisms of the argument for being one-sided, for expressing the opinion of the implied author, or for lacking statistical support. A few candidates criticised the author for being biased against universities and/or parents.
- (d) Most candidates realised that if university education is free to the students, it has to be paid for from taxation, but some thought it would mean that professors and other staff would not be paid, which they rightly rejected as unrealistic. Many candidates argued that if students did not have to pay for their tuition, large numbers of young people who lacked the academic potential for success would enter university and either waste their time or gain a degree which they did not deserve. Not many differentiated between admission based on wealth and admission based on ability and academic achievement. A good proportion of candidates on this occasion used strands of reasoning, and thereby scored 4 or 5 marks, but many still seemed to have written down unrelated ideas as they occurred to them, which limited their mark to 3. A few candidates were awarded 0, 1 or 2 marks, mostly because they explored the practicalities and implications of making university education free to students, instead of arguing for or against it.

THINKING SKILLS

Paper 9694/22 Critical Thinking
--

Key messages

- Many candidates are still giving a summary of the content of a source before moving on to actually answering the question. This is unnecessary and wastes time.
- A minority of candidates still have problem with timing, spending far too long on the short answer parts of **Question 1** and subsequently running out of time in **Question 3**. Such candidates often produce lengthier answers to the three-mark questions than to the six-mark questions.

General comments

Candidates found **Question 2(a)** and **(b)** easier than last summer. Candidates seemed to respond well to the issues raised by the questions and were able to tackle them effectively. The trend in recent papers of most candidates understanding the nature of the examination continued. Such candidates realised that expressing opinions about the issues raised or showing further knowledge was not the key focus. However, there is still a minority of candidates who seem unprepared and struggle to reach a total mark in double figures.

Comments on specific questions

Question 1

A number of candidates made the unwarranted assumption that Sources D and E were talking about the same thing and that Grade C mortar also took a long time to set. However, an encouraging number of candidates saw the implications in the information in Source E for the quality of the wall if Mr Jones had to rush.

- (a) Candidates found this question rather difficult and often just repeated the information in the source. The relevance of vested interest was subtle. Candidates could not be credited for simply saying that Jones had a vested interest to lie. At first sight the reverse is true, and he says things against himself. However, this could be a cunning plan to conceal the real lie that the problems were only cosmetic. Very few candidates developed their answer along these lines.
- (b) A significant number of candidates saw the possible relevance that Smith's daughter may have damaged the wall sufficiently by driving into it for it to fall over as a result of the tremor. Some candidates did not consider the importance of the expression 'possible relevance' and dismissed the information because there was no proof that she had collided with the wall. Some bizarre lines of thought emerged in some of the answers. It was suggested that the tremor may have been a result of Smith's daughter driving in to the garage wall. Also, that the damage to the buildings may have been a result of Smith's daughter driving in to them rather than the earth tremor.
- (c) Again, a significant number of candidates saw that even Grade C mortar was guaranteed for 10 years so using it did not have any relevance to the wall falling down shortly after it had been built. As noted above, some candidates mistakenly thought that Grade C mortar also took longer to set. Other candidates struggled with the expression 'recently' and thought that this might apply even after 10 years.

- (d) Clear conclusions were reached in most cases. Some candidates confused the question of Jones's general skills as a workman with his workmanship on this particular job. This often led them to blaming Mr Smith for hurrying Jones. Whilst this is an explanation for why his work was not up to its usual standard it does not alter the fact that his workmanship was poor in this case. Few candidates explored the point that even if his workmanship was a bit poor (as Jones himself admitted) it might not have been so poor that it caused the wall to fall down. As noted above, many candidates made good use of Source E to make inferences about why the wall might have fallen down as a result of Jones having to rush.

Question 2

- (a) (i) Although this question had a deceptively simple answer, it was quite a technical question involving the identification of the assumption (all music has lower frequency bass notes) that needed to be made in order to draw the conclusion. This question discriminated well with a tendency for more able candidates to answer it correctly.
- (a) (ii) Candidates found it difficult to express themselves in answering this question. They needed to say either:
- It did not undermine it at all, because a point about maintaining sound quality did not alter the fact that vinyl has the capacity to produce better sound quality.
- or
- It did undermine it because, in practice, nobody would be able to keep up this maintenance regime so they would not produce a purer musical sound after a few plays.
- (b) Candidates needed to initially see that this statement would undermine the logic of the argument in Source C, as it would run counter to what one would expect if the reasoning was correct. However, it only counters it if one makes the assumption that the quality of the viewing experience is the prime reason why people go to the cinema. Many candidates went straight for the 'second' part of the answer. Some candidates thought relative numbers going to the cinema and viewing on their phone was significant but the sheer rise in numbers going to the cinema was sufficient information to undermine the argument in the source. A minority of candidates seemed to misunderstand the nature of the question and simply said the statement could not be true as Source C says that the numbers are declining.
- (c) This was an easy question but many candidates did not gain marks because:
- Points made were aspects of the same thing
 - Points made were about recording method
 - Points made were about the subjective judgement of the music. This was not relevant to the quality of recorded sound as such – one can judge sound quality even if one doesn't like or understand the music.
- (d) Many candidates used the sources effectively in this question, usually to argue that critics were old-fashioned and prejudiced. A number made an effective distinction between being old-fashioned and being irrationally prejudiced. Others thought a prejudice could be rational which opened up a line of thought that probably went beyond a six-mark question. A minority of candidates thought the statement was arguing that digital recordings/supporters of digital recordings were old-fashioned and prejudiced suggesting they need to read questions more carefully.

Question 3

- (a) This question discriminated well and there were an encouraging number of two-mark answers.
- (b) Many candidates identified three intermediate conclusions and the vast majority identified at least one intermediate conclusion.
- (c) Many candidates gained two marks, usually by identifying the inconsistency between the pessimistic view of science/humanity expressed in the second paragraph with the optimistic view expressed in the final paragraph. Fewer candidates were able to identify the straw man flaw in paragraph 2. A minority of candidates still challenged the statements made rather than evaluating the reasoning in the passage, but the majority of candidates are now understanding the nature of the exercise in this question.
- (d) Candidates did less well in this question than in past papers. There was a tendency to argue that countries should come to a global agreement, whereas the proposition was whether or not they could. Others argued that individuals and groups could do a lot to combat global warming, but this rather took the focus of global agreements which implied formal agreements between countries. Candidates who successfully argued against the proposition tended to do so via a reference to past examples of global agreements. Candidates who argued for the proposition avoided the above problems and came out with a number of good lines of reasoning re incompatibility of interests and economic development.

THINKING SKILLS

Paper 9694/23 Critical Thinking
--

Key messages

- Many candidates are still giving a summary of the content of a source before moving on to actually answering the question. This is unnecessary and wastes time.
- A minority of candidates still have problem with timing, spending far too long on the short answer parts of **Question 1** and subsequently running out of time in **Question 3**. Such candidates often produce lengthier answers to the three-mark questions than to the six-mark questions.

General comments

Candidates seemed to respond well to the issues raised by the questions and were able to tackle them effectively. The trend in recent papers of most candidates understanding the nature of the examination continued. Such candidates realised that expressing opinions about the issues raised or showing further knowledge was not the key focus. However, there is still a minority of candidates who seem unprepared and struggle to reach a total mark in double figures.

Comments on specific questions

Question 1

- (a) Most candidates explored the implication of the incriminating remarks by Singh about his car's performance. More able candidates then modified this by suggesting he might have been boasting or that he would not want to have damaged his new car.
- (b) A significant number of candidates did not focus on relevance as regards responsibility for the accident. Others focussed on reliability rather than relevance, with the expertise of the engineer being most often cited. Very few candidates made the point that we do not know that the road Singh was on would fall in to the category of having 'considerable' traffic. We know traffic had increased but this does not necessarily mean the traffic is considerable.
- (c) Many candidates struggled with this question. Vested interest was relevant but not in the standard way that candidates attempted. We already know that he has a vested interest in the matter so to be told he is a councillor doesn't add anything to this. It does however reinforce the view that farmers have an influence on the council and that they will be sympathetic to the farming interest.
- (d) Most candidates came to a clear conclusion that Singh was responsible. Many took advantage of the opportunity to explore a plausible alternative that the farmers/council were to blame. More able candidates pointed out that Singh was local and should have known about the mud hazard. They also pointed out that signs about farm vehicles made no explicit reference to mud.

Question 2

- (a) Many candidates identified the conclusion and saw that eliminating drugs would not be a sufficient condition for ensuring long-term protection for health. However, very few related this to the threats to health that arise from the competitive nature of sport, of which performance-enhancing drugs are only an aspect.
- (b) Only a minority of candidates understood this question. A large number explored the implications for disabled athletes competing with able-bodied athletes, which was not the focus of the question. Others did not focus on the issue of fairness. The crucial point was that the use of 'cyborg' type technology might mean a Paralympics of the future might just be a question of which athlete had the better technology.
- (c) A significant number of candidates misunderstood the statistics and thought that Honduras had the best results because they had more national records. Those who interpreted the statistics correctly were able to produce a variety of explanations. As usual with this sort of question, candidates need to guard against producing different versions of essentially the same explanation.
- (d) A significant number of candidates had difficulty interpreting this statement and often ended up with reasoning which supported the opposite of the position they had chosen. Typically, this was a case of declaring their support for the proposition and then using arguments that suggested the ban on drugs in sport should continue. More able candidates were able to make effective use of the sources to construct good answers, especially supporting the proposition.

Question 3

- (a) Most candidates got one- or two-mark answers.
- (b) Many scored three marks.
- (c) There were significantly more four- and five-mark answers to this question than in past papers. Popular points of evaluation were that noise reduction was not a sufficient condition to conclude that a floor covering was superior (though candidates rarely used the technical expression), the assumption that jobs in wooden floor manufacture have not outweighed the loss of jobs in the carpet industry, the *ad hominem* reference to "old hippies" and the inconsistency as regards the reasoning about wool and synthetic materials in paragraphs 5 and 6. Candidates were less successful in evaluating the reasoning in paragraph 3, where there was a tendency to suggest that one had to assume no other injuries were possible (as opposed to injuries from hard surfaces). This was too broad an assumption. As in other papers, the minority of candidates who misunderstand the nature of the exercise in **Question 3(c)** were unable to take advantage of these evaluation opportunities and could not receive any credit for simply challenging the propositions in the passage.
- (d) The proposition produced good arguments both for and against and there were many four- and five-mark answers. Opponents of the proposition tended, unsurprisingly, to focus on the non-biodegradable nature of plastic, often with considerable technical detail. Candidates need to be wary of using too much factual content in constructing arguments in this question. More credit is given for two or three separate strands of reasoning rather than one strand of reasoning with a lot of empirical support. Credit is given for examples but they must be used sparingly.

THINKING SKILLS

Paper 9694/31 Problem Analysis and Solution
--

Key messages

The search for short-cuts and general rules when tackling complex problems under timed conditions is a key part of successful problem-solving. But any short-cut that is found must be scrutinised and tested where possible – or you may build the entire analysis of a problem on a misapprehension. This is particularly the case when extending the understanding of a relationship from a small, manageable number of cases, to a general case (as in **Question 3**). Success at this process is at the heart of good problem-solving, and is as dependent on sceptical self-checking, as it is on the speculative creation of ‘formulae’.

General comments

Working: erroneous working should have a neat line though it, so that it can be read by an Examiner. Relevant working should not be crossed out. Even where arithmetic is poor, partial credit can often be obtained if calculations are shown and numbers have text to indicate what they are. This is particularly the case where a problem involves a search for an optimal solution – any attempt is better than no attempt (and will normally be awarded partial credit), and any attempt to improve on a first attempt is better still.

Explanations: candidates must be alert to the distinction between showing that a certain solution is possible, and showing that it is inevitable. The latter is harder, and requires explicit consideration of the alternatives – and comment on why they should be discarded. Such explanations may involve the consideration of a number of steps, and it is easy to get lost in the midst of them. To counteract this, candidates should endeavour to signpost their explanations as clearly as possible: labels such as “Vouchers A & E applied to all books”, “Ratio of colours needed”, “Why ‘0’ is not possible” are a good idea.

Comments on specific questions

Question 1

This question required candidates to work with discrete units (two table sizes, with different ‘per person’ costs) to find an optimal solution to a collection of closely-related problems. The problems are ‘Diophantine’ and so can not be easily analysed algebraically. The variety of charging options and the number of ways they could be combined invited an orderly review of the search space, in order to be sure that proposed answers were optimal. On the whole, the question was answered well by most candidates.

- (a) Many candidates offered a correct answer without working – and were awarded full marks. There was no need to confirm that solutions were exhaustive, because the question asked specifically for the ‘the two possible combinations’. A small number of candidates gave answers which involved the right arithmetic, but did not clearly state the number of round and square tables – and were only awarded 1 mark as a result.
- (b) The restriction on the number of tables required candidates to revisit the combinations considered in (a) – once again there was no need for an exhaustive search, since the fact that there is only one answer is implicit in the way the question is asked (“Which combination....”).
- (c) 44 people can be precisely catered for, and this question required candidates to calculate how much this would cost. This was successfully accomplished by most.

- (d) Success at this question depended on an organised approach to combining the different options. Most candidates attempted to list the costs for the 44 people and 46 people, paying per person, per table and for the whole room. Those who laid this out clearly tended to reach the correct pair of 'cheapest prices' and find the difference between them. Many candidates launched into their survey of the prices with little care for the record they left, and then misjudged their own analysis.
- (e) Success at this question depended on an appreciation of how it differed from (d). The requirement that Alan use the same payment method limited the combinations, and made a comprehensive survey easier to accomplish. A substantial number of candidates used incorrect numbers of tables and people, and reached incorrect conclusions – the discipline of laying out a solution for an audience to understand did seem to correlate with correct analysis and solution here.

Question 2

Success at this question depended on the confident calculation of percentage changes and a tactical choice of how the collection of vouchers could be applied most effectively.

- (a) Most candidates appreciated that the maximum discount would be achieved by applying a voucher to three \$60 books. And many concluded that the greatest discount for the four was found by applying voucher L. A few candidates incorrectly assumed that no two books could have the same price (and so began with books priced \$60, \$59 and \$58).
- (b) This question introduced the tension between voucher S and voucher L: in particular, the fact that voucher S offers a maximal discount when the total cost of the books is greatest, whereas voucher L offers a maximal discount when the lowest priced book is greatest. \$32 was a common, sub-optimal answer.
- (c) (i) This question introduced the tactics of how the vouchers are to be combined – and depended on careful and confident calculation of percentage change. The question asked for a demonstration that Brock is correct, and hence clear working was vital. Most candidates left a solution which clearly established which vouchers were applied to which books. From this it became clear that a substantial minority had misinterpreted the application of voucher E, interpreting "any single book" as "every single book". Such distinctions in the use of qualifying words are vital in these questions.
- (ii) This question did not ask for supporting working – and hence was technically easier – but asked for a pair for which the order would **never** matter. Many candidates offered a solution which mirrored what they had done in (i), showing a pair which would not offer different discounts when applied to particular combinations of books.
- (d) This final question required candidates to reflect on the lessons learned in (b) and (c), and then seek an optimal way of combining them. Without a strategy, the search space was daunting. The most common solution involved buying all the books together, and applying the vouchers in the order SALE. Care was needed in applying voucher L here, since the price of the cheapest book was affected by the application of S and A. Many candidates did not perform the necessary technical calculations correctly, as a result of trying to find short-cuts. A small number of candidates appreciated that voucher L was used most effectively on the three most expensive books, and attempted this combination – which gave the optimal answer.

Question 3

This question required candidates to consider how simple flag designs could be cut from striped material – the relationship between the number of flags and the length of material was not obvious, and most candidates either based their working on an oversimplified relationship or struggled to visualise the problem at all.

- (a) This question was designed to invite candidates to diagram the key relationship, and for this to act as a springboard for the more speculative questions that followed. Very few candidates offered a diagram, and very few offered correct answers as a result. The most common error involved a hasty attempt to express the relationship as a linear rule (e.g. $15 + 30n$, where 15 represents the unused material at the beginning, and 30 represents the length of a flag). Such oversimplification of the question tended to yield incorrect answers to (i) and (ii). Those who drew a sketch of the appropriate number of stripes tended to get both questions right.

- (b) Those who misjudged (a) were clearly in a weak position for (b), which depended on generalising the relationship. For those who had characterised the relationship correctly, care was needed in considering the final section – and partial credit was awarded for appropriate working, or neighbouring solutions.
- (c) This question was more amenable to the ‘simplifications’ that many candidates were tempted to do – the amount of white material defined how many flags could be made, and a calculation of how many white stripes would occur in 3 m of material yielded the correct answer. This question was accomplished successfully by many candidates.
- (d) The problem posed here was independent of the challenges of (a), (b) and (c) – and explicitly asked for a diagram. Tactically, it involved considering the ratio of white to blue material in the first 2 m (1 : 1) and then combining this with the amount of blue and white material used by Andy to create a new ratio. There were a number of easy mistakes that could be made here – overlooking the need to combine the flags with Andy’s, overlooking the length restriction, reversing the ratio – and very few candidates offered an appropriate answer. Most candidates did offer answers which showed their intended lengths and colours clearly.
- (e) This question involved calculating the appropriate lengths of material (for 4 flags with and without joins) and then costing the two options. Those who had struggled with (a) and (b) tended to persist with inappropriate modelling of the relationship here. Very few candidates offered correct prices for both options.
- (f) This final optimisation question depended on tactical application of the relationships discovered in the earlier parts of the question: as a result, very few candidates were able to tackle the problem successfully.

Question 4

This question involved the relationship between numbers in a grid, and their sums and squares. It depended on playing with the numbers, and their squares, in the context of the game ‘Square Deal’. Almost all candidates were able to engage with the problems posed, and a significant number tackled this question first.

- (a) The only errors here tended to be in the arithmetic.
- (b) Full marks here depended on a careful explanation of why 2 and 4 were **necessarily** the tiles that Gordon selected – not just an explanation of why they **could have been** the tiles he selected. The most common answer which did not suffice involved calculations showing that 2 and 4 would deliver scores of 158 and 74.
- (c) This question required a reflection on tactics, before submitting a viable grid: many candidates appreciated that a sum of 25 in both a row and a column was desirable; only a few appreciated that the constituent numbers summing to 25 affected the score, and that the number of 8s and 7s should be maximised. Most candidates were awarded marks for offering a viable grid, which added to their credit for (c)(i), even if little strategic thinking had been deployed.
- (d) As with (b), success at this question depended on clear explanation of choices, along with careful and explicit consideration of the alternatives. Candidates chose a number of different ways of describing where the two tiles should be placed – and were credited with marks as long as the descriptions were unambiguous. Descriptions such as ‘bottom left in Russell’s grid’ were credited (a unique identifier, as long as only the two empty spaces are considered) but such brevity is not encouraged. The implication of why the 1 was to be placed in the 2nd row down and 2nd column of Russell’s grid was achieved by most candidates. The implications for the placing of the 5 required much greater care, and many candidates did not look beyond its immediate placement (“Gordon will get 0 if the 5 is placed in either square”). Those candidates who listed the four remaining tiles – a key constituent of a correct analysis of where the 5 was to be placed – tended to score 3 or 4 marks here.

THINKING SKILLS

Paper 9694/32 Problem Analysis and Solution
--

Key messages

To gain full marks, candidates will often need to be clear whether the context uses whole numbers or a continuous range, especially when dealing with percentages, and to be aware if they are faced with a quantized model of a continuous process. They should adapt rounding to the scenario as appropriate, e.g. up to complete hours rather than apply a rigid rule in all cases.

General comments

The examiners are looking for demonstration that the candidates can both do the analysis and provide at least one solution, so the working that produces the answer is vital, and, to gain full credit, is not something to be crossed out when the final answer has been produced (especially when it is given and the task is to explain how it is obtained). Discarded working should have a neat line through it. Relevant working should not be crossed out. Even where arithmetic is poor, partial credit can often be obtained if calculations are shown and numbers have text to indicate what they are. Where 4 or 5 marks are available there will be more to do, but partial credit is available for good analysis, even where no solution is offered.

Many candidates could save time and reduce errors by a more efficient approach to numeric calculations. For example, a ten per cent reduction on a total could be simply calculated as $0.9(a+b+c)$, rather than a more laborious method such as $((a-10a/100) + (b-10b/100) + (c-10c/100))$.

A few candidates appear to have run out of time, but almost all attempted most of each question. About a third gave no response to the last parts of **Questions 2, 3 and 4**.

Those who put two vertical lines in \$ should ensure that their handwriting could not be misread as \$1.

Comments on specific questions

Question 1

This question's focus was on the most efficient use of adults for supervising children in a nursery. The numbers were chosen so that different limitations and approaches needed to be applied for the cases which superficially appear to be repeating the same question.

- (a) Some candidates tried to maximize the use of rooms rather than optimize the use of people. Better responses avoided trying to use linear algebra when it was inappropriate, but there were many erroneous responses.
- (b) For full marks, candidates needed to show that their answer was achievable, not just an inspired guess.
- (c) More marks could have been obtained if more candidates had checked that their proposed answer was feasible.
- (d) The best responses offered a systematic approach to show the best possible configuration. Candidates who gave the same answer using the same method three times might usefully have wondered if they might have missed some detail.

- (e) Better responses offered a concise description of the pertinent reason that one was best, and avoided producing an essay with a range of plausible but unsupported suggestions.

Question 2

A model using whole numbers in order of size was presented, and questions were about the best or worst cases.

- (a) Most candidates were able to find the extreme value.
- (b) (i) For full marks, candidates need to be clear if they are showing that something cannot be done, by using a contradiction, rather than merely recording their lack of success in finding something.
- (ii) Candidates needed to be careful to distinguish the information given in the text from the hypothetical parts of earlier parts of the question which no longer applied. Here, the assumption in (a) about everyone having at least \$1 no longer applied.
- (iii) Only a few candidates took account of the implications of the ordered numbers, e.g. all of the individual items in the middle 80% necessarily being not less than the highest in the bottom 10%.
- (c) Some candidates saw that the limit would be the most uniform distribution, but only a few took proper account of the constraint that whole numbers were required.
- (d) Only the best responses noted the constraint that the middle 8 10% could not have less than the lowest 10%.

Question 3

This question considered the stocking and restocking of a shop. Data was provided on the stock levels, from which changes needed to be calculated and considered.

- (a) There were many arithmetic errors in answers, often caused by incorrect and unnecessary conversion of time to seconds.
- (b) When candidates make a slip, they can often work out that they have made a mistake using a reality check: is this about right in this context?
- (c) Many candidates appeared to have addressed a much simpler question of their own; the fact that three marks were available strongly suggests that a non-trivial response was required.
- (d) Only the best responses gained full marks, by taking account of the different speeds of the two people in handling the different sorts.
- (e) Some candidates considered the week's sales to be between Monday and Friday morning, which were shown in the table, rather than Monday and the same time on the following Monday, the data for which was written in the text but not in the table.

Question 4

This question explored the options available, costs, and finding the way for a group to do all they wanted, subject to various practical constraints.

- (a) A very common slip was to forget that there are $(m - n + 1)$ numbers from m to n inclusive.
- (b) Some candidates did not remember that there are 60, not 100, minutes in an hour.
- (c) Some indication that the answer given is as good as possible was needed for full marks.
- (d) Attention to detail was often lacking, e.g. some included the two people in the count of those in the queue behind them.

- (e) This required a calculation of what was needed, deduction of what they already had, and then to determine the most cost-effective combination to purchase enough extra. Only a few candidates scored full marks.
- (f) Better responses found the most tightly-constrained options first, making it easier to cut down the possible options for the rest, but a few responses were internally inconsistent, e.g. having a person at two places at one time.

THINKING SKILLS

Paper 9694/33 Problem Analysis and Solution
--

Key messages

To gain full marks, candidates will often need to be clear whether the context uses whole numbers or a continuous range, especially when dealing with percentages, and to be aware if they are faced with a quantized model of a continuous process. They should adapt rounding to the scenario as appropriate, e.g. up to complete hours rather than apply a rigid rule in all cases.

General comments

The examiners are looking for demonstration that the candidates can both do the analysis and provide at least one solution, so the working that produces the answer is vital, and, to gain full credit, is not something to be crossed out when the final answer has been produced (especially when it is given and the task is to explain how it is obtained). Discarded working should have a neat line through it. Relevant working should not be crossed out. Even where arithmetic is poor, partial credit can often be obtained if calculations are shown and numbers have text to indicate what they are. Where 4 or 5 marks are available there will be more to do, but partial credit is available for good analysis, even where no solution is offered.

Many candidates could save time and reduce errors by a more efficient approach to numeric calculations. For example, a ten per cent reduction on a total could be simply calculated as $0.9(a+b+c)$, rather than a more laborious method such as $((a-10a/100) + (b-10b/100) + (c-10c/100))$.

A few candidates appear to have run out of time, but almost all attempted most of each question. About a third gave no response to the last parts of **Questions 2, 3 and 4**.

Those who put two vertical lines in \$ should ensure that their handwriting could not be misread as \$1.

Comments on specific questions

Question 1

This question's focus was on the most efficient use of adults for supervising children in a nursery. The numbers were chosen so that different limitations and approaches needed to be applied for the cases which superficially appear to be repeating the same question.

- (a) Some candidates tried to maximize the use of rooms rather than optimize the use of people. Better responses avoided trying to use linear algebra when it was inappropriate, but there were many erroneous responses.
- (b) For full marks, candidates needed to show that their answer was achievable, not just an inspired guess.
- (c) More marks could have been obtained if more candidates had checked that their proposed answer was feasible.
- (d) The best responses offered a systematic approach to show the best possible configuration. Candidates who gave the same answer using the same method three times might usefully have wondered if they might have missed some detail.

- (e) Better responses offered a concise description of the pertinent reason that one was best, and avoided producing an essay with a range of plausible but unsupported suggestions.

Question 2

A model using whole numbers in order of size was presented, and questions were about the best or worst cases.

- (a) Most candidates were able to find the extreme value.
- (b) (i) For full marks, candidates need to be clear if they are showing that something cannot be done, by using a contradiction, rather than merely recording their lack of success in finding something.
- (ii) Candidates needed to be careful to distinguish the information given in the text from the hypothetical parts of earlier parts of the question which no longer applied. Here, the assumption in (a) about everyone having at least \$1 no longer applied.
- (iii) Only a few candidates took account of the implications of the ordered numbers, e.g. all of the individual items in the middle 80% necessarily being not less than the highest in the bottom 10%.
- (c) Some candidates saw that the limit would be the most uniform distribution, but only a few took proper account of the constraint that whole numbers were required.
- (d) Only the best responses noted the constraint that the middle 8 10% could not have less than the lowest 10%.

Question 3

This question considered the stocking and restocking of a shop. Data was provided on the stock levels, from which changes needed to be calculated and considered.

- (a) There were many arithmetic errors in answers, often caused by incorrect and unnecessary conversion of time to seconds.
- (b) When candidates make a slip, they can often work out that they have made a mistake using a reality check: is this about right in this context?
- (c) Many candidates appeared to have addressed a much simpler question of their own; the fact that three marks were available strongly suggests that a non-trivial response was required.
- (d) Only the best responses gained full marks, by taking account of the different speeds of the two people in handling the different sorts.
- (e) Some candidates considered the week's sales to be between Monday and Friday morning, which were shown in the table, rather than Monday and the same time on the following Monday, the data for which was written in the text but not in the table.

Question 4

This question explored the options available, costs, and finding the way for a group to do all they wanted, subject to various practical constraints.

- (a) A very common slip was to forget that there are $(m - n + 1)$ numbers from m to n inclusive.
- (b) Some candidates did not remember that there are 60, not 100, minutes in an hour.
- (c) Some indication that the answer given is as good as possible was needed for full marks.
- (d) Attention to detail was often lacking, e.g. some included the two people in the count of those in the queue behind them.

- (e) This required a calculation of what was needed, deduction of what they already had, and then to determine the most cost-effective combination to purchase enough extra. Only a few candidates scored full marks.
- (f) Better responses found the most tightly-constrained options first, making it easier to cut down the possible options for the rest, but a few responses were internally inconsistent, e.g. having a person at two places at one time.

THINKING SKILLS

Paper 9694/41 Applied Reasoning
--

Key messages

- **Question 2** is a relatively straightforward question for those who understand the task. All that is needed is a bullet-pointed list of statements copied from the text and labelled as MC, IC or CA.
- In **Question 3** many candidates are still gaining no 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

It was pleasing to note that there were fewer candidates writing answers whose length does not reflect the mark allocation and most candidates appeared to have enough time to finish the paper.

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

Comments on specific questions

Question 1

Most marks could be gained from a thorough discussion of the idea that correlation does not imply a causal link and the majority of candidates were able to gain some credit for this approach, although very few explored it in enough depth to achieve more than two marks. Many candidates noted the invalid comparison between populations, either in terms of geographical area or marital status of respondents. All the other points on the mark scheme were seen, but rarely. Very few candidates scored full marks but most got at least one. Vague references to generalisation of data from one province were not credited. A number of candidates did not gain any marks; often this was because they merely questioned the source of the data or because their answer explained how the data supported the claim.

Question 2

Performance on this question was better than in previous series. Most candidates now know what they are expected to do. Hence it discriminated well between prepared candidates. A high proportion could correctly identify the main conclusion and those who had been taught how to recognise ICs and CAs often scored between four and six marks. Some did not know what was required at all but, unfortunately, there were those who seemed able to identify the various elements, but did not label them and so could not be given any credit.

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 had been taught some critical evaluation were able to readily identify the slippery slope. A smaller number could describe some problems with the use of definitions and the *post hoc* nature of the Second World War example. Some candidates were even able to identify some of the unstated assumptions, and descriptions of other points on the mark scheme were seen occasionally. Some candidates appeared to be able to see where there was a weakness but, without understanding of the task, they merely copied out the 'weak' section rather than explaining what was weak about it. Many candidates still do not understand the technical term 'assumption' and label any opinion-based claim in the text as such. If it has been stated in the text it *cannot* be an unstated assumption.

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 the countries of the world should unite as one. Many found they could engage with this topic, and responses contained more of the candidates' own thinking and ideas than in previous series. Candidates also appeared to be able to arrange their ideas into strands of reasoning more competently, and there was less sequential treatment of documents than in previous series; hence marks for the Structure and Quality skill areas were a little higher than usual. 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. For the first time in this paper, Documents 3 and 4 contained multiple viewpoints. This gave an opportunity to candidates to look for corroborative or conflicting statements within as well as between documents. However, many candidates treated these documents as a holistic single viewpoint, which did not make much sense and detracted from the effectiveness of their argument. 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

- **Question 2** is a relatively straightforward question for those who understand the task. All that is needed is a bullet-pointed list of statements copied from the text and labelled as MC, IC or CA.
- In **Question 3** many candidates are still gaining no 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 candidates appeared to have enough time to finish the paper but some did not. Often, those who did not have time to complete **Question 4** had spent a long time on **Question 3**, achieving marks anyway, when marks would have been achieved more easily on **Question 4**. It was pleasing to note that there were fewer candidates writing answers whose length does not reflect the mark allocation.

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

Comments on specific questions

Question 1

Both parts of the question invited candidates to question the strength of a newspaper claim based on the information given. Many were able to engage with this and achieve one or two marks for each section. It was rare to award three marks in part **(a)** and many candidates scored 0 in either or both sections. Marks were awarded for all ideas listed on the mark scheme. In part **(b)** all points on the mark scheme were seen, the most common being a reference to the re-evaluation of existing evidence rather than the emergence of new. A number of candidates did not gain any marks; often this was because they criticised the source of the information, rather than the way the information was used to support the claims given.

Question 2

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

Question 3

The better-prepared candidates attempted to evaluate the passage, but many are still listing a series of counter-arguments to points in the passage. Those candidates who did attempt to apply their evaluation skills were often able to gain some marks. The most frequently credited weakness identified was the slippery slope in paragraph 3. Other weaknesses commonly identified were the inappropriate example in paragraph 2 and the causal flaw in paragraph 3. Others were rarely discussed to a creditworthy level but a promising minority of candidates did successfully spot and articulate 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 people who use their freedom of expression to cause harm should be severely punished. The most successful candidates directed their reasoning towards all three aspects of the conclusion: freedom of expression, causing harm and severe punishment. However, some candidates struggled with the multi-layered conclusion that was given. Many adapted the conclusion to one that was easier to construct a case towards, sometimes focussing solely on freedom of expression, or arguing towards a vague conclusion about the punishment matching the severity of the harm caused. While this approach allowed them to access most of the marks, there were not able to access all the marks in the quality of argument skill if they did not reason towards the given conclusion.

The majority of candidates scored below twelve marks, because their reasons rarely went beyond what was written in the documents, and the information given in the documents was used uncritically. A minority of candidates were able to produce a thoughtful and reasoned case that achieved higher marks. It was pleasing to see some candidates attempt to structure their arguments using strands of reasoning and intermediate conclusions, and only a very few neglected to state any conclusion at all. A small number of candidates simply described the contribution made by each document to the debate. A few Centres seem to have acted upon advice given in previous reports 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

- **Question 2** is a relatively straightforward question for those who understand the task. All that is needed is a bullet-pointed list of statements copied from the text and labelled as MC, IC or CA.
- In **Question 3** many candidates are still gaining no 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 candidates appeared to have enough time to finish the paper but some did not. Often, those who did not have time to complete **Question 4** had spent a long time on **Question 3**, achieving marks anyway, when marks would have been achieved more easily on **Question 4**. It was pleasing to note that there were fewer candidates writing answers whose length does not reflect the mark allocation.

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

Comments on specific questions

Question 1

Both parts of the question invited candidates to question the strength of a newspaper claim based on the information given. Many were able to engage with this and achieve one or two marks for each section. It was rare to award three marks in part **(a)** and many candidates scored 0 in either or both sections. Marks were awarded for all ideas listed on the mark scheme. In part **(b)** all points on the mark scheme were seen, the most common being a reference to the re-evaluation of existing evidence rather than the emergence of new. A number of candidates did not gain any marks; often this was because they criticised the source of the information, rather than the way the information was used to support the claims given.

Question 2

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

Question 3

The better-prepared candidates attempted to evaluate the passage, but many are still listing a series of counter-arguments to points in the passage. Those candidates who did attempt to apply their evaluation skills were often able to gain some marks. The most frequently credited weakness identified was the slippery slope in paragraph 3. Other weaknesses commonly identified were the inappropriate example in paragraph 2 and the causal flaw in paragraph 3. Others were rarely discussed to a creditworthy level but a promising minority of candidates did successfully spot and articulate 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 people who use their freedom of expression to cause harm should be severely punished. The most successful candidates directed their reasoning towards all three aspects of the conclusion: freedom of expression, causing harm and severe punishment. However, some candidates struggled with the multi-layered conclusion that was given. Many adapted the conclusion to one that was easier to construct a case towards, sometimes focussing solely on freedom of expression, or arguing towards a vague conclusion about the punishment matching the severity of the harm caused. While this approach allowed them to access most of the marks, there were not able to access all the marks in the quality of argument skill if they did not reason towards the given conclusion.

The majority of candidates scored below twelve marks, because their reasons rarely went beyond what was written in the documents, and the information given in the documents was used uncritically. A minority of candidates were able to produce a thoughtful and reasoned case that achieved higher marks. It was pleasing to see some candidates attempt to structure their arguments using strands of reasoning and intermediate conclusions, and only a very few neglected to state any conclusion at all. A small number of candidates simply described the contribution made by each document to the debate. A few Centres seem to have acted upon advice given in previous reports 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.