

# PSYCHOLOGY

---

Paper 9698/11

Core Studies 1

## General comments

As with all papers, there was a spread of questions on different aspects of the studies, such as background, apparatus, procedure, results and conclusions. In **Section A**, the candidates' knowledge of background (7(a), 14(a)), procedure (1(a)(b), 4(a), 5(a), 6(a), 8(a)), results (2(a)(b), 3(b), 6(b), 7(b), 14(b)), conclusions (8(b), 11(b)) and evaluation (1(b), 4(b), 5(b)) was good. However, some parts of **Section A** of this paper presented particular challenges to some candidates. In general many candidates could improve by having a better general understanding of the context of the study, for example to tackle **Question 8(a)**, and of the procedure in some cases (e.g. **Questions 15(a) and 15(b)**). To improve performance still further, candidates would benefit from a more effective grasp of methodology in psychology so that they can see how the study illustrates these principles, for example to be able to improve their answers to **Questions 2(a) and 15(b)**.

Some candidates offered good responses in **Section B**, writing essays that were relevant and focused on evaluation rather than description. Many candidates could, however, improve their answers by illustrating their evaluative points with examples from the content of the chosen study.

## Comments on specific questions

### **Section A**

#### **Question 1**

- (a) Most candidates who scored here discussed the crimes committed or truths and lies. However, a small number gave excellent answers referring to being known to the police or their proficiency in English which facilitated very good answers to part (b).
- (b) As mentioned above, good answers were given in relation to understanding of English and being known to the police, for example indicating that non-fluent speakers might have paused for longer while working out how to express a truth, confounding the results or that those who were not known to the police might have been more frightened by the interview process and ummed more even when truth-telling just because it was unfamiliar.

#### **Question 2**

- (a) A number of candidates appeared to be unaware how a mean is worked out, so were therefore unable to earn full marks on this question (although they could earn partial marks for identifying that the average calculated was a mean). Other candidates showed a sophisticated understanding, including an awareness that the false and true story totals had to be divided by different values.
- (b) Most candidates were able to score marks here, reporting the direction of the difference accurately and often giving accurate values for the results. A small number offered only conclusions rather than results.

#### **Question 3**

- (a) Many candidates were able to score some credit here. The number of candidates reported conclusions rather than results here. Another error was to report gender differences in the AS/HFA group (which could not have been found, since the group consisted only of male participants).

- (b) Most candidates understood that adults with autism spectrum disorders had lower social intelligence (than the control participants). However, there was often confusion about non-social intelligence, with a large number of candidates mistakenly claiming that Baron-Cohen et al. concluded that people with ASD had higher non-social intelligence than the control participants. Candidates need to be aware that these two variables are *not* related. Another problem was that some candidates producing good answers in respect of poor social intelligence sometimes said nothing at all about non-social intelligence, so were only responding to half of the question. One final occasional error was to report nonexistent differences in gender recognition by the AS/HFA group.

#### Question 4

- (a) All three possible tests were well described by some candidates. Even the pupillary reflex test, which candidates found difficult to describe as the word 'dilate' was sometimes misunderstood, produced many answers that were sufficiently clear for full marks. The most common error was to describe the blink test or to produce a muddled answer which combined a description of the visual pursuit and blink tests. Another occasional error was to name two tests rather than to 'describe' any (as required by the question).
- (b) A common error here was to repeat or rephrase the stem of the question. Nevertheless, there were many good answers referring to the inability to test the IV of visual experience on visually guided behaviour if the passive kitten could not see.

#### Question 5

- (a) There were many full mark answers to this question. The most common error was to cite the purpose given to participants as testing 'learning and memory' without mentioning punishment. A small number of candidates also mistakenly reported that the participants were told it was a study about authority or obedience, so could not earn marks.
- (b) This question part was generally quite well answered, although some candidates gave answers that were too general (e.g. 'to reduce demand characteristics') which needed to be more closely related to the study in order to gain higher marks.

#### Question 6

- (a) Many candidates scored full marks on this question part, with accurate reporting of a range of the administrative routines. Where errors were made they were most commonly to include other aspects of the prison set-up, for example describing what happened when the prisoners first arrived, such as being given a prisoners' uniform.
- (b) There were many good answers here, generally referring to the increase in duration of the head count. Where candidates did not earn marks their answers typically mistakenly focused on pathology of power in general.

#### Question 7

- (a) This question was very well answered, with candidates often giving elaborated answers earning full marks. However, a proportion of candidates misreported effect, suggesting that time to respond decreases as the size of the group increases.
- (b) The majority of candidates were able to answer this question part correctly, although a few gave muddled answers as for part (a). Some candidates successfully argued 'yes' effectively, on the basis of comments such as 'I didn't help because he should, he's a man' or that even in big groups people did consider individual responsibility.

#### Question 8

- (a) Very few candidates were able to identify both 'groupness' and 'fairness'. Indeed, even the 'groupness' mark was often earned serendipitously.

- (b) In spite of the weaker responses to part (a) of this question, answers here were rather better. Many candidates were able to offer well explained examples of the consequences for society from negative attitudes to the out-group in terms of prejudice or discrimination, for example the reluctance to help others in need when they are from other social groups. One fairly common error was to report the results of Tajfel's study which could not earn marks as the question required the consequences for society.

### Question 9

Although this question was fairly well answered, there were few full mark answers. This was largely because candidates were responding by 'identifying' rather than 'describing' their two chosen categories. Some candidates scored no marks at all because they described imitative rather than non-imitative behaviours.

### Question 10

- (a) A proportion of candidates were unable to earn full marks because they only referred to two of the three experimental groups that were required by the question. Very few candidates offered accurate data, although note they could earn full marks without any data if they offered all three experimental groups in the correct order.
- (b) This question was well answered with many candidates knowing both that further data was collected by observation and that the results were the same. Some candidates offered explanations of the data or conclusions, which were not creditworthy as the question asked for results.

### Question 11

- (a) Many candidates partially misunderstood the question, and instead of outlining two dreams which both contained horizontal and vertical eye movements, they offered two dreams, one with horizontal and one with vertical eye movements.
- (b) Even those candidates who had been unable to score full marks in part (a) of this question were generally able to score at least one mark here, and this question part was generally quite well answered. Where errors were made, these were often because the candidate had misread the question and instead of responding to 'how these eye movements were explained', They gave answers describing 'how this eye movement data was collected', for example describing the role of the EEG.

### Question 12

- (a) Although this question was generally fairly well answered, there was one common mistake. Candidates tended to reword the question rather than answering it. Many candidates simply repeated that the task was that 'the participants described a route through London', without giving any additional information.
- (b) This question part was also well answered, with almost all candidates being able to identify a comparison task and many giving an effective explanation of its purpose. The majority of candidates chose to explain another experimental task rather than the baseline task, although either were appropriate. A small number of candidates mistakenly referred to the film scenes/frames task, which was incorrect as this comparison was not made to the routes task.

### Question 13

- (a) Answers to this question part were variable, including basic answers centring around the idea that Eve Black appeared only when she wanted to, and more sophisticated ones describing the way in which Eve Black could imitate Eve White.
- (b) This question part was well answered, with many candidates accurately reporting the difference in IQ and a small proportion describing the results of the Rorschach test for both personalities. A few candidates did not earn marks because they described differences that were not based on a test, such as evidence from the letter.

**Question 14**

- (a) This question part was fairly well answered with many candidates being able to offer at least a partial definition of systemising. Some candidates did not earn marks here, for example those who described systemising as 'male-type thinking/behaviour'.
- (b) This question was quite well answered. Candidates were more likely to score full marks by providing relative comparisons for both males/female and for physical sciences/humanities than for providing accurate data.

**Question 15**

- (a) Although a small number of candidates mistakenly described the experimental group rather than the control group, the majority were able to give some appropriate information. Few were aware that the participants were age- and sex-matched, which was unexpected as candidates are likely to have discussed the differences in mirror uses by males and females.
- (b) This question part was not as well answered as part (a). Although a proportion of candidates were aware that the control group were recruited from personal contacts, surprisingly few of these were able to identify this as an example of opportunity sampling method.

**Section B**

**Question 16**

The most popular choice here was Rosenhan although there were also good answers based on Demattè et al. and a small number using Langlois et al. The quality of responses was varied, with many candidates able to identify basic aspects of validity such as the need for controls or the importance of the generalisability of the test situation for ecological validity. Better essays were able to offer examples of these aspects from their chosen study and the best essays expanded on these basic ideas with a range of other strengths and weaknesses in terms of validity. Although some candidates were confused about the difference between reliability and validity, many were clear on this conceptual point. However, some candidates merely evaluated the study in general, earning marks only by coincidence, or just described the study, earning no marks at all.

**Question 17**

In answering this question, candidates' choices were spread over all three studies, with the majority choosing Bandura et al. or Freud. Few answers were clearly focused on the developmental approach, thus there were not many high scoring responses. Nevertheless, many candidates were able to make relevant points within their answers, so scored some marks. Candidates would benefit from considering the strengths and weakness of each of the approaches in general, to enable them to focus on the key ideas so that they would be able to use any available study to generate examples.

# PSYCHOLOGY

---

Paper 9698/12

Core Studies 1

## Key Messages

- Candidates should answer questions in the order in which they are presented on the question paper. Question 16 or 17 from **Section B** could be done before **Section A** questions, but **Section A** questions should not be done out of order.
- The writing of some candidates is difficult to read and all candidates are encouraged to write legibly. Candidates should not use ink which bleeds through the page as this makes the reverse side difficult to read.
- Candidates should provide answers that equate to mark allocation, so an answer worth 2 marks should be short and an answer worth 10 marks should be correspondingly longer. **Section B** questions are not short-answer.
- For a **Section A** 2-mark answer that has the command ‘describe’, candidates should ensure they provide enough detail to score both marks, rather than a partial, very brief or vague answer.
- Candidates should read all parts of a question, **(a)** and **(b)** in **Section A** before beginning to write an answer to ensure that the answers to both question parts are not the same.
- Candidates should answer *both* parts within the same question and where there are two parts they should ensure both parts are answered.
- Candidates should look to quote psychological knowledge wherever possible. Anecdotal answers will never achieve top marks.
- Candidates should always seek to evaluate using psychological methods, approaches, issues and debates as appear in the syllabus rather than with general evaluation points.

## General Comments

The detail included in this report will help candidates improve their examination technique as well as giving comments on how specific questions should be answered.

Many candidates re-wrote the question, or the introduction to the question, as their answer, which is not necessary and should be discouraged. Candidates are advised to read both parts of the question. This was most evident in **Question 11**. When a question has two components, such as *identify* and *explain*, candidates must answer both of these components, marks will be allocated to both components. See **Questions 7(b), 10(b) and 11(a)**.

## Comments on Specific Questions

### Question 1

- (a) Many candidates identified these two behaviours, gaze aversion and pauses, and scored full marks. Some candidates identified one of these behaviours, for one mark, often giving blinking, incorrectly, as the other (behaviours analysed using frequencies were not required by the question).
- (b) Candidates who correctly identified pauses and gaze aversion in (a) often scored full marks for this question part; others scored some credit for correctly writing about pauses.

## Question 2

- (a) Correct answers included knowledge about *another person* being lost or visiting a shopping mall themselves for example. Because the point of the study was to create a false memory, candidates who said that the participant had been lost in a shopping mall in the past could not be credited.
- (b) Loftus and Pickrel believe that false memories evolve when old memories become linked to (or confounded with) the new suggestion, so that the memory that the false event was a mere suggestion deteriorates and it becomes combined with (or part of) real memories. Many candidates understood this and scored full marks. However, many did not and many guesses were attempted.

## Question 3

- (a) When revising the eyes test Baron-Cohen et al. hoped that the revised test would remove the ceiling effect, so that the (normal) scores were not all clustered at the top of the range. This correct answer would score full marks and many candidates achieved this. A few candidates went on to provide further explanation only to realise that what they had written here was the answer to question **part (b)**. Candidates should always read both question parts before beginning their answer.
- (b) Some candidates re-wrote their **part (a)** answer. The revised eyes test was better because it could now discriminate smaller (individual) differences between the different groups of participants rather than just detecting extreme differences in performance.

## Question 4

- (a) The passive kitten could move in a circle (forwards or backwards, left or right), up and down (bend its legs), or it could turn around in either direction where it was standing. Credit was given for each of these, and many candidates scored full marks. Some candidates described other things such as the visual cliff, which could not be credited.
- (b) This question part was misunderstood by many candidates, most of which described the movements of the two kittens (active kitten moves forwards so the passive kitten also moves forwards). This question is asking about the difference in kittens' *experience*, not the effect of that difference on their behaviour. The perfect answer, achieved by some candidates, was that the active kitten felt movement and saw visual changes at the same time, whereas the passive kitten saw visual changes without simultaneous movement.

## Question 5

- (a) A large number of candidates said that the participant saw the stooge receiving the shocks. This is incorrect, because the stooge/learner was in a different room from the participant/teacher. Some candidates said that the participant heard the cries and screams of the learner but the question did not ask what the participant *heard*. There are three correct answers which are (quoting from the core study) "Upon depressing a switch: a pilot light corresponding to each switch is illuminated in bright red; an electric blue light, labelled 'voltage energizer', flashes; the dial on the voltage meter swings to the right."
- (b) Many candidates answered this question incorrectly because they failed to answer the question set, why the events that were *seen* (i.e. what the participant saw) were important. Any other answer, such as what was heard, could not be credited. Candidates needed to explain that the events were important so that the participant would believe that they were responsible for shocking the learner; that they would believe that the shocks were increasing; that they would be aware of the shock intensity.

### Question 6

- (a) The correct answer to this question was any member of the prison staff or an ex-prisoner. Many candidates over-complicated their answers writing incorrectly about the President, State Governor or about people who actually designed a prison such as an architect.
- (b) Many candidates repeated the question, ‘it was to make it realistic’. What was important was that it created and maintained the state and psychological imprisonment. Another reason why it needed to be realistic was that it was important to allow it to be generalisable or applicable to real prisons. A few candidates scored no marks when writing that “it makes it reliable and valid” without further explanation or understanding.

### Question 7

- (a) This question required candidates to describe the critical area of the subway car, not what happened in the critical area of the subway car. A description of the latter scored no marks. Many candidates wrote that there was a pole in the middle of the subway car which was also correct. A number of candidates described where the observers were sitting. This was incorrect because the observers sat in the adjacent and not the critical area of the subway car.
- (b) This question part required two components: firstly the number of people leaving the critical area and secondly a brief explanation of why those in the critical area might have done so. For the first part 34 people left the critical area (on 21/103 trials). For the explanation any logical reason why people might have left the critical area would receive credit. For example, because leaving reduces emotional arousal, it avoids any cost of helping such as effort or causing any physical harm to people.

### Question 8

- (a) Common errors were to describe validity as ‘when something is valid’ or ‘when something is reliable and valid’, which were not creditable responses. Credit was given to those candidates who referred to a type of validity, such as ecological validity. Candidates who scored maximum marks were those who stated that validity is the extent to which a test (or measure) tests what it claims to.
- (b) The Tajfel study was valid because it measured what it claimed to. It was claimed that people discriminate on the basis of very little such as being put into a ‘minimal group’. The boys in the study were unaware that their grouping was random and so any stated differences between them, such as when they opted for maximum difference must have been caused by their allocation into the meaningless groups. Whilst a few candidates understood this and scored maximum marks, many candidates did not.

### Question 9

- (a) Many candidates wrote very strong answers to this question, but many candidates did not know why children’s pre-existing levels of aggression were rated.
- (b) A common incorrect response was that inter-rater reliability increases or ensures high reliability. Candidates needed to explain that it is simply a measure of how much agreement (or disagreement) there is between two independent observers. If the data gathered by the observers is similar there is high inter-rater reliability as measured by the application of a test of correlation (score closer to 1) and if there is little agreement in the data gathered there will be low inter-rater reliability with a correlation closer to 0.

### Question 10

- (a) A number of candidates rewrote the words of the introduction to the question, which scored no marks, stating that little Hans believed that all animate objects had a penis. Candidates needed to explain that Hans expected his little sister to have a penis and expected her ‘small penis’ to grow; also for his mother to have a (big) penis and he expected his own penis to grow when he got older. Whilst many candidates scored full marks, some did not largely through misunderstanding the question.

- (b) This question required candidates to both outline **and** explain one dream or fantasy. Many candidates failed to address both components of the question. Some candidates described the event of his mother powdering his widdler, this could not be credited because it wasn't a dream or fantasy.

#### Question 11

- (a) Most candidates were credited for naming a guideline, but some did not go on to describe that guideline. A number of candidates were vague when stating that one guideline is "harm" or "consent" when at this level the full term should be known which are 'protection from physical (or mental) harm' and 'informed consent'.
- (b) Candidates are always advised to read both parts of the question, i.e. (a) and (b) before starting to write their answer. In this instance some candidates could not apply what they had written in **part (a)** to **part (b)** and some candidates give a different guideline in **part (b)** from that given in **part (a)**, despite **part (b)** question stating '*how this* ethical guideline was followed'. A common error was that many candidates gave the guideline of informed consent, stating (and falsely assuming) that the participants had given full written informed consent. Participants did complete a 'general health' questionnaire, but they did **not** give informed consent. Demattè et al. state "All participants were naïve to the purpose of the study at the beginning".

#### Question 12

- (a) Most candidates gained some credit for explaining that the pseudo-patients were nervous because they believed they would be exposed as frauds, and some of these candidates gained further credit for expanding, saying that this would be embarrassing because the study would come to an end. Other candidates said that it was because most of them had never visited a psychiatric ward before or that they had genuine fears about what might happen to them. For these answers marks were determined by the amount of detail provided, with brief or muddled answers gaining limited credit.
- (b) A 'suggest' question requires candidates to think about how something could be applied to a given situation. A suggestion is not something that can be learned beforehand and it requires candidates to think for themselves during the examination. Some candidates re-wrote their **part (a)** answer, others made a partial suggestion. Strong responses extended one point to clarify their answer and others provided two brief points, both resulting in full marks. One way in which the nervousness was good was because real patients would be nervous too, and so it adds to realism; one way in which the nervousness was bad was that it may have contributed to the appearance of abnormality and so show symptoms even though they weren't part of the claim they were hearing voices.

#### Question 13

- (a) This question required specific results rather than general comments or a comparison. For example some candidates wrote simply that Eve White had a higher IQ than Eve Black, which is too vague. Some candidates wrote a correct result for one Eve but an incorrect result of the other and scored just one mark. Many candidates scored full marks.
- (b) Candidates could answer this question from the viewpoint of either Eve White or Eve Black. Any appropriate comment was acceptable. For example some candidates wrote that Eve White was a nervous person (1 mark) and so she didn't concentrate on the test (2 marks). Some candidates wrote that Eve Black had a mischievous and playful personality (1 mark) and so she didn't take the test seriously, explaining why her score was lower (1 mark). Some candidates scored no marks when they wrote that a personality could be either regressive or repressive, but neither of these would explain why an IQ score was higher or lower.

#### Question 14

- (a) Most candidates scored full marks when answering this question.
- (b) Like **question 12(b)** this was a 'suggest' question. Any appropriate answer could be credited, such as increasing the interest of girls, or giving additional 'scaffolding' or support to show how to apply the system-related tasks.

### Question 15

- (a) Many candidates scored some credit for this question because they stated that a self-report questionnaire was used. Very few candidates scored full marks because they did not provide any detail about the questionnaire used to measure long mirror sessions.
- (b) Many candidates gave either no answer or an incorrect answer for this question. Some candidates did score full marks. The correct answer was that although BDD patients were shown to experience more distress than controls before both short and long sessions, it wasn't possible to compare the effect of resisting these sessions. Credit was also given for answers such as 'they didn't want to cause further distress'.

### Question 16

Many candidates scored only a few marks because they did not approach the question in the correct way, they either described the study (which could not be credited), or evaluated it without reference to the physiological approach. Some candidates started by explaining what the physiological approach is but then went on to provide general evaluation. A number of candidates who did approach the question correctly failed to score marks because of technique errors. These included giving either weaknesses or strengths but not both, gaining limited marks. Relevant points could be made about the scientific equipment used to measure the physiological processes being studied; that physiology is the same in all humans and so a relatively smaller sample could be used with no cultural differences (for example, adrenaline has the same physiological effect on all people). On the negative side, a physiological approach is reductionist and applied to the Schachter and Singer study they show that it is the cognitive interpretation of what is happening that is important.

### Question 17

A number of exam technique errors were evident. Some candidates do not understand the term 'discuss' which must include both strengths and weaknesses to be a discussion. Any answer that provides only strengths or only weaknesses will gain limited credit, however strong the answer is. For a perfect answer there can be some imbalance in the length of a strength or a weakness, but if there is at least two of each, answers can access the top mark band. As was the case with question 16 many candidates gave specific evaluation points that were not related to snapshot studies. For example, if a participant is required for only a few minutes, then larger samples can be gathered. Many candidates appear to believe that *all* snapshot studies are inexpensive. This is incorrect, many snapshot studies are very expensive to conduct. There is also the belief that they are 'easy to do' and they do not 'waste the time of the experimenter'. Comments like this should be avoided, as they are generalised and inaccurate.

# PSYCHOLOGY

---

Paper 9698/13

Core Studies 1

## General comments

As with all papers, there was a spread of questions on different aspects of the studies, such as background, apparatus, procedure, results, conclusions and evaluation. In **Section A**, the candidates' knowledge of background (1(a)(b), 8), apparatus (**Question 4(a)**), procedure (3,10(b)), results (4(b),5(b)), conclusions (2(b)) and evaluation (7,11) was good. However, some parts of **Section A** of this paper presented particular challenges to some candidates. In general many candidates could improve by making sure that they are responding to the demands of the question. For example, for both parts of **Question 7** an effective answer needed to be applied to Bandura et al's study, which many candidates did not do even though they were able to give adequate partial answers which they probably could have readily related to the context of the study. This also applies to **Question 8(b)**.

It was pleasing to see that some candidates offered good responses in **Section B**, writing essays that were relevant and focused on evaluation rather than description, especially in response to **Question 16**. Some candidates would, however, benefit from remembering that for answers to earn marks they require evaluation illustrated with examples from the content of the chosen study rather than just description of that study.

## Comments on specific questions

### **Section A**

#### **Question 1**

- (a) Candidates were generally able to correctly identify that the suspect's name was not used, and many were also able to gain the second mark for 'address'. However, marks were often lost for non-specific answers such as 'their personal details'.
- (b) Many candidates correctly stated that this information was not included because 'easy' truths were not comparable to 'difficult' lies, and the aim was to investigate high-stakes lying (and this mark was often earned, irrespective of the score in part (a)). Some candidates, however, demonstrated an excellent understanding of the study, explaining that simple truths would place very much less demand on cognitive load than more complex truths making them inappropriate comparisons to the high-stakes lies.

#### **Question 2**

- (a) The most frequent answers centred around explanations of the small sample and/or the lack of generalisability. More proficient answers linked the ideas to individual differences and illustrated this with the variety of responses from participants.
- (b) This was generally well answered, with candidates typically describing the claim that false memories really do exist and expanding on this with reference to eyewitnesses and courts of law.

#### **Question 3**

This question generated a wide spread of answers, some responses provided more than enough information on two groups but others were only just able to identify a single group accurately. Few candidates spent time providing more information than was necessary by describing more than two groups.

#### Question 4

- (a) Although many candidates were able to describe the gondola, there were some common errors, such as that the kitten could not move its head or could not move its legs. It could in fact move both, within the limits of the gondola (but was unable to control its own movement).
- (b) There were some good answers here, with appropriate comparisons between the passive and active kittens (or with reference to the passive kittens after spending time in the light). However, many candidates lost marks by referring to the visual cliff test rather than to visually-guided paw placements.

#### Question 5

- (a) Although some candidates were able to report correctly that self-evaluations became more negative for both prisoners and guards, it was a very common error for candidates to report that the self-evaluations of the guards became more positive.
- (b) This question was generally well answered, and often produced more accurate answers than part (a). However, some candidates simply described the guards' negative behaviour rather than their *verbal* behaviour. Other candidates were unable to gain marks because they described pathology of power (i.e. described the cause of the change in their behaviour).

#### Question 6

- (a) Where candidates earned marks here it was generally for answers which hit upon the idea that the conditions were different from earlier studies on which diffusion of responsibility was based, such as that the passengers were in a confined space so were confronted with the emergency.
- (b) Again, answers which earned marks here largely did so through a good knowledge of psychology and the making of sensible suggestions rather than a sound knowledge of the study itself. Nevertheless, this generated some appropriate responses.

#### Question 7

- (a) Very many answers here earned only one mark because although they made comments of varying detail, they needed to be linked to the Bandura et al. study to answer the question set and gain full marks. Where links were made, they were generally those indicated on the mark scheme, although in addition, candidates made reference to the spreading of aggressive children evenly across levels of the independent variable (i.e. the matching process).
- (b) Again in this question part there were plenty of partial answers that lacked a link to the study to gain full marks. However, a small number of candidates identified a possible ethical disadvantage: that the participants might have suffered distress from seeing the aggressive model or long-term effects from the aggressive modelling, which would not have been imposed in a case study. A good range of methodological issues were explained, such as the problem with low ecological validity in the experimental setting of the observation room compared to watching a child's aggressive behaviour at home, the problem of not knowing about the pre-exposure of each child to aggression before experiment (compared to the detail that would be known in a case study) and the idea that Bandura et al. could not study long term effects of exposure to aggressive models in their experiment as they only observed for 20 minutes (but in a case study you could follow a child for longer).

#### Question 8

- (a) There were some good answers here which fully explained the role of the ego in balancing the id and the superego. Weaker answers simply suggested that the ego was 'the self' which is not a sufficient description alone.
- (b) Those candidates who mentioned the id and superego in part (a) were then often able to produce excellent answers relating the id and superego to Hans's problems, leading to full marks, for example explaining that the ego balanced Hans's desire for his mother (id) and knowing he should not (superego).

**Question 9**

- (a) Most candidates were able to report the finding that the infants looked for longer at the attractive babies (so preferred them) although a common error was simply to say that ‘there was a difference in attractiveness’, without indicating whether this meant that they preferred attractive or unattractive babies. Some candidates, however, did not expand on the simple statement of looking for longer at attractive babies, so could not gain full marks, whilst more successful candidates offered elaboration such as that this difference was present regardless of the babies’ gender, or that this provides further evidence that attractiveness is based on invariant (perhaps universal) cues.
- (b) Responses here were mixed. Where candidates knew that the infants could not tell the difference between male and female babies they were often also able to expand on this to identify that this result was the same as for adults (who could not discriminate either) or very accurately reported that there was a slight preference for same-sex babies, but that this was non-significant. However, other candidates incorrectly believed that (unlike adults) infants could detect gender in other babies.

**Question 10**

- (a) Some candidates were able to identify the immediate task, i.e. that they children were asked a verbal question and some elaborated this to describe that they were given three alternatives, of ‘good boy’, ‘bad boy’ or ‘just okay’. However, other candidates mistakenly reported that the children were asked to describe the story immediately, which was not the case.
- (b) Many candidates were able to accurately report that the scale had 7 pictures, but others simply repeated the content of the question (that it was a scale of faces) whilst better candidates described the key differences, such that this was a visual task rather than a verbal one (so was easier to understand) or that it was more informative as there were 7 choices not just 3.

**Question 11**

- (a) Although many candidates correctly gave raised heart rate, fewer reported palpitations or tremor, and other less successful candidates mistakenly reported the false side effects that were told to the EPI-MIS group.
- (b) A number of candidates understood that this as because the experimental conditions (of apparent emotional state) had not ‘worked’, but the explanations were often muddled or unelaborated. More successful candidates explained the problem clearly or added that the effect would have been to confounded the results.

**Question 12**

- (a) This question was generally well answered, with candidates reporting a range of appropriate answers, most often that it provided information about the participant’s brain waves, indicating their stage of sleep and whether they were dreaming.
- (b) Candidates typically gave thoughtful answers, for example observing that it would give the participant freedom of movement, and continuing this to say that they would therefore sleep more normally (so increasing ecological validity). Other candidates raised appropriate practical issues, such as it reduced the risk of unplugging the equipment or detaching the electrodes.

**Question 13**

- (a) Although almost all candidates were able to identify an appropriate experimental condition, few related this to the results of the speech output baseline task.
- (b) Reference to the differences in speech output was often absent from answers or incorrect. Only a few candidates were able to describe the difference, that the participants would have paused to think about their route (for example), so would have spoken less in the experimental conditions than in the control (baseline) repetitive speech output.

#### Question 14

- (a) This question was well answered, in a variety of ways, most often with reference to Demattè et al. using more smells, so being able to make more relevant comparisons, for example because people do not smell of banana or some people like bananas but others do not. Other appropriate comments on validity included saying that the presentation of the smells was brief in the Demattè et al. study, so it was less likely to affect mood, or that the use of the olfactometer/dilutions ensured the delivery of the smells was always equivalent.
- (b) Although this question was quite well answered, candidates were a little less successful here than in part (a) at relating their answer to the study, and the marks gained were limited. For example, responses often gave little more than a generic response of 'it was repeated measures so the participants might have become fatigued', 'it lacked ecological validity', 'It could have just been halo dumping' or 'it only used women'. Better answers expanded on these such ideas, for example saying 'It was repeated measures so the participants might have become fatigued by the repeated smells and been less able to tell them apart', 'It lacked ecological validity because people really experience attraction in social situations/not with their head placed on a chin rest', 'It could have just been halo dumping i.e. transferring feelings from one to another i.e. smell to face' or 'It only used women and men might be differently affected by smells'.

#### Question 15

- (a) Many candidates were able to identify at least one domain, although there were few full mark answers here.
- (b) Some candidates were able to describe the structure of INPUT – OPERATION – OUTPUT and a small number were able to add some elaboration to this comment.

#### Section B

#### Question 16

Rosenhan and Thigpen and Cleckley were the most popular choices here, although some essays on Billington et al. were also seen. The quality of responses was varied although many were able to identify basic points about ecological validity, such as generalisability beyond the setting, for example the extent to which Rosenhan's research, being set in a range of hospitals, would have been relevant elsewhere and whether Eve's behaviour could have been a product of the psychiatrist's office. However, many answers typically earned low band marks, as they gave only referred to one way in which the study did (or did not) have ecological validity or gave relatively generic answers rather than relating the points being made to the study. A small number of responses contained only description of the study so did not answer the question and could not be credited. It is important that candidates are equipped with a suite of tools to answer these questions, specifically an understanding of the range of issues that can be asked about plus a good knowledge of the study so that they apply the ideas to the study and use this to find examples to illustrate their essay. In addition, even those essays by candidates who were able to supply some appropriate points relating to ecological validity and apply them to the study often focused exclusively, or almost exclusively, on either strengths or weaknesses, so gained limited credit.

#### Question 17

The most popular choice here was Milgram, although some candidates also tackled the essay using Tajfel or Veale and Riley. The quality of responses was also varied, ranging from answers which seemed unaware of the meaning of quantitative data or were unable to offer any evaluation of the study in this respect, offering instead only description (and therefore scoring 0 marks) to those which demonstrated a clear understanding of a variety of strengths and weaknesses of collecting quantitative data in psychology and illustrating these with reference to their chosen study. Milgram answers were the most likely to be able to offer both strengths and weaknesses and a range of thoughtful applications to the study were seen.

# PSYCHOLOGY

---

Paper 9698/21

Core Studies 2

## Key Messages

### **Section A**

#### **Question 1**

It is important that candidates are made aware of the issues in psychology as some were unable to give a detailed definition of the case study as a research method in **part (a)**. Candidates should suggest a simple alternative to the original study in **part (b)** and give clear details of the procedure followed. In addition, candidates need to be aware of the features of each method as many gave a description of a study with a large sample size which is not a case study. Extended evaluative points that make direct reference to the alternative idea are necessary in **part (c)** to achieve full marks.

#### **Question 2**

It is important that candidates practice writing these types of questions. Some did not structure their responses appropriately and could not achieve full marks. For example, if the question asks for strengths and weaknesses then four points must be made (two strengths and two weaknesses). Candidates must refer to the named study in their responses to achieve higher marks.

### **Section B**

Candidates must write more extended responses in both **part (b)** and **part (c)** of the essay as many gave accurate responses that lacked depth. Evidence must be given in **part (c)** to achieve higher marks.

## General comments

The marks achieved by candidates sitting this examination covered the whole spectrum of the mark scheme. Many did provide good answers which showed that they were very well prepared and did consistently refer to the evidence in order to achieve high marks.

Time management for this paper was good for most candidates and most attempted all questions that were required.

A small minority of candidates answered both questions in the **Section B** essay. When a candidate did this they were awarded the mark for the better of the two questions (**Question 3** or **Question 4**). These candidates usually achieved very poorly.

Candidates need to cover the entire syllabus so that they can respond to the questions in **Section A** where there is no choice of question. In addition to this, candidates must include evidence in the **part (c)** of their **Section B** essays to achieve higher marks. **Question 3** was the slightly more popular choice of question.

## Comments on Specific Questions

### **Section A**

#### **Question 1**

- (a) Most candidates could achieve some marks for this question. Popular choices included describing the single or small sample of participant, in-depth nature and longitudinal aspects of case studies.

A few candidates achieved full marks by going on to describe that case studies can often show change or development over time and the researchers can become quite close to the participant(s) and therefore gain a lot of insight into their behaviour.

- (b) Most candidates were able to describe a procedure that used the case study method and was measuring REM and dreaming in some way. Quite a few ‘mirrored’ ideas from the original Dement and Kleitman study ensuring that the ‘how’ and ‘what’ elements of the procedure were covered and often included details of sample. There were still a minority of candidates presenting experimental work with large pools of participants so could only gain partial credit.

Some candidates wrote out results of their study which did not achieve any marks. Popular ideas included doing the study in the home of the participant and using the same machinery as Dement and Kleitman or more complex machinery available today such as an MRI or CAT scan. Most candidates focused their response on briefly describing how the alternative study would be carried out and what data would be collected. Many did not include the other details required, such as where the study would take place and who the participants would be. In addition, the candidates need to ensure the ‘what’ and ‘how’ for the procedure are very clear. For this particular study, the candidate did need to give details of the equipment used if this was relevant to their procedure.

A small minority of candidates evaluated their idea in this question and received no credit for this as this is the correct response to **Question 1(c)**.

- (c) The vast majority of candidates achieved marks in this question by providing some evaluative points. Most gave both methodological and practical issues in their response.

Many discussed issues about the ecological validity of their study, practical issues with using equipment at the home of the participant, lack of generalisability of the small sample group and ethical issues that might come about as a result of a long and/or in-depth nature of their study. Most also contextualised their responses referring directly to their original idea.

A few gave well developed points that achieved very high marks as they referred directly to their alternative idea. Some only briefly identified issues and did not refer back to the context of their own study.

## Question 2

- (a) The majority of candidates achieved full marks for their answer to this question. They referred to the fact that quantitative data is numerical data. They often extended their response by referring to the ability to use statistics, graphs or the possibility of making comparisons with quantitative data. Some did not achieve full marks as they gave a very brief response.

- (b) Many candidates received marks for this question by giving a finding from the study. Popular responses were to describe the finding that the boys rewarded more points to their own group than to the out-group. They also chose the maximum difference in awarding points between the in- and out-group. A lot of candidates simply identified what quantitative data had been collected rather than focusing on a finding.

- (c) The vast majority of candidates achieved some marks in this section. Most were able to describe one strength and one weakness of collecting quantitative data and were able to give an example from the Tajfel study. Many referred to the data being easy to compare and also that statistical tests can be used with quantitative data. For the weaknesses most referred to the lack of depth in the data.

There are still many candidates not attempting the ‘plural nature’ of these types of questions. Candidates need to describe two strengths and two weaknesses to achieve higher marks. Many did attempt to do this, but found it difficult to describe an appropriate second weakness (e.g. validity and decreased usefulness).

- (d) Many candidates answered appropriately for this question and were able to achieve higher marks. Many candidates could identify at least one positive or negative issue linked to applying Tajfel’s study to everyday life. Issues such as the generalisability of the sample of boys and lack of ecological validity were discussed effectively.

A number did describe the usefulness of the study and gained limited credit. Some candidates just discussed how the study explained categorisation and prejudice without any reference to the application to everyday life and gained limited credit.

## **Section B**

### **Question 3**

- (a) The vast majority of candidates achieved credit for this question and were aware that reductionism referred to the simplification of something more complex. Many achieved full marks and gave a clear definition.
- (b) Many candidates attempted to answer the question and wrote about how the data were collected in each study. Some just described the procedure and achieved some marks for this description. The Bandura et al. study produced the best answers and many were able to give detailed descriptions of the observations undertaken during the study. For the Baron-Cohen et al. study, most very briefly described or just named the eyes task and often did not then go onto describe the other methods of data collection (e.g. IQ test, AQ test and gender recognition task). The Maguire et al. study produced the lowest marks for candidates. Many incorrectly thought an MRI scanner was used and Centres must ensure they teach the correct Maguire et al. study. Some mentioned the PET scanner and gave details of how it worked, and strong responses included details of the tasks the taxi drivers were asked to do. Very few mentioned that it was the hippocampus that was scanned.
- (c) Many candidates identified one problem with carrying out reductionist research and many gave an example from a study to back up their point. Some were able to identify two or three ideas but need to give sufficient evidence to back up their points. The most popular problem discussed was the oversimplification of the findings.

### **Question 4**

- (a) There were many good responses to this question and many candidates achieved full marks by giving a clear and often quite detailed definition of the cognitive approach. Weaker were brief or muddled.
- (b) Candidates achieved poorly on this question. Many could give general descriptions of some of the procedures of the studies but could not describe the cognitive processes investigated in each study. Some understood the requirements of the question very well and gave a focused response that achieved higher marks.
- (c) For this question, candidates need to identify and discuss three advantages psychologists have when they investigate cognitive processes with clear reference to a core study for each point. Many were able to describe one or two advantages such as the applications of the findings of studies and how they can help to improve everyday life and issues with laboratory experiments which make up the majority of cognitive research. Many did not link their responses to a study and achieved fewer marks. Also, many discussed problems psychologists have when they investigate the cognitive processes which was not required by the question.

# PSYCHOLOGY

---

Paper 9698/22

Core Studies 2

## Key Messages

### **Section A**

#### **Question 1**

Candidates should be aware of the requirements of each question in the exam. For example, if asked to describe they should not include evaluative comments. Many did describe the features of self report in part (a) but they also evaluated this method which did not receive any additional credit. Candidates should suggest a simple alternative as a self report to the original study in part (b) and give clear details of the procedure followed, including the sampling method used. Extended evaluative points that make direct reference to the alternative idea are necessary in part (c) to achieve full marks.

#### **Question 2**

It is important that candidates practice writing these types of questions. Some did not structure their responses appropriately and could not achieve full marks. For example, if the question asks for strengths and weaknesses then four points must be made (two strengths and two weaknesses). Candidates must refer to the named study in their responses to achieve higher marks.

### **Section B**

Candidates must write more extended responses in both part (b) and part (c) of the essay as many gave accurate responses that lacked depth. Evidence must be given in part (c) to achieve higher marks.

## General comments

The marks achieved by candidates sitting this examination covered the whole spectrum of the mark scheme. Many provided good answers which showed that they were very well prepared and did consistently refer to the evidence in order to achieve high marks.

Time management for this paper was good for most candidates and most attempted all questions that were required.

A small minority of candidates answered both questions in the **Section B** essay. When a candidate did this they were awarded the mark for the better of the two questions (**Question 3** or **Question 4**). These candidates usually achieved very poorly.

Candidates need to cover the entire syllabus so that they can respond to the questions in **Section A** where there is no choice of question. In addition to this, candidates must include evidence in the part (c) of their **Section B** essays to achieve higher marks. **Question 3** and **Question 4** were equally popular with candidates.

## **Comments on Specific Questions**

### **Section A**

#### **Question 1**

- (a) Most candidates achieved higher marks for this question. Candidates were able to describe elements of the self-report method, very few candidates couldn't provide a creditable answer. More common answers included questionnaire and interviews, qualitative and quantitative data and open and closed questions. Some candidates focused too much on evaluation issues or giving examples of questions which were not creditable.
- (b) Most candidates were able to describe a procedure that did use the self report method and investigated intergroup categorisation. Some candidates did not do either a self report or focus on the intergroup categorisation and gained limited marks. Common ideas were to divide the participants into groups on a random basis and then ask questions about what they thought of the other group or what they might be willing to give either their group or the other group.

There were a number of candidates who wrote a fully replicable procedure, but some candidates did not refer to how the sample would be obtained or where the study was done.

A very small minority of candidates evaluated their idea in this question and received no credit for this as this is the correct response to **Question 1(c)**.

- (c) The vast majority of candidates achieved marks in this question by providing some evaluative points.

Many discussed issues about the ethics of studying participants who are children as well as the deception used in their study to allocate the participants into groups on a random basis. Other popular points included evaluation of the type of data (qualitative and/or quantitative), ecological validity and generalisability of the sample.

Many gave well developed points that achieved very high marks as they referred directly to their alternative idea. A significant number of candidates only briefly identified issues and did not refer back to the context of their own study.

#### **Question 2**

- (a) The majority of candidates achieved some credit for their response to this question. Candidates tended to give basic answers for just discussing children. The best answers understood that the developmental approach concerns the changes in behaviour overtime.
- (b) This question was not well answered. Many candidates did not distinguish between the behaviour of the younger and older children but simply referred to children in general, consequently achieving no marks. A significant minority gave findings from the wrong study. Some were able to achieve marks and the most popular response was to refer to the fact that the older children recalled more accurate details of the story compared to the younger children. A few could then relate this finding to the developmental approach which shows our memory does improve as we age, and achieved full marks.
- (c) The vast majority of candidates achieved some marks in this section. Most were able to describe at least one strength and one weakness of the developmental approach. Many referred to the approach being useful and also that children can answer truthfully or without demand characteristics as they are unaware they are in a psychology study. For the weaknesses some referred to the ethical problems of studying children as well as the issues surrounding concentration and understanding problems in very young children.

However, a significant number of candidates referred to the strengths and weaknesses of the Nelson study and achieved fewer marks as a result. Many also referred to the issues of generalising from children to adults which is not relevant to this approach as it focusses on the age of the children in the study and does not require to be generalised to adults.

Many candidates attempted the ‘plural nature’ of these types of questions, some describing two strengths and two weaknesses to achieve the higher marks.

- (d) Many achieved well on this question by discussing the use of children and giving some examples from the Nelson study. Popular points included ethical issues, studying development at different ages, problems with access to children and/or ease of access to children in nurseries and schools, and communication problems.

Better answers discussed fewer issues in more detail while the lower mark answers tended to mention many issues but in less depth and often without reference to the Nelson study.

## Section B

### Question 3

- (a) Many candidates used an example of a psychometric test rather than stating what is meant by the term and achieved limited credit. Some candidates however did explain the term accurately as the mathematical measure of the mind and achieved full marks for their response.
- (b) This question was well answered by many candidates. Most achieved high marks and many achieved full marks. The most common answers for the Baron-Cohen et al. study were the eyes test and the IQ test although the eyes test was not described in sufficient depth. The most common answers for the Thigpen and Cleckley study was the 100 hours of interviews, IQ test, ink blot and EEG. For the Billington et al. study common answers were the EQ and SQ-R. Candidates tended to struggle the most with this study in terms of outlining how the data were collected.
- (c) Most candidates could identify advantages of using psychometric tests although most found it difficult to identify three advantages. Popular points included usefulness, comparable results and reliability of the tests. Some candidates referred back to the core studies as evidence to back up their points. However, many did not do this and therefore achieved lower marks. Many referred to non-psychometric tests in their responses as examples and these were not creditworthy e.g. the eyes test from the Baron-Cohen et al. study.

### Question 4

- (a) Most candidates were able to achieve limited credit by referring to ecological validity or usefulness in their response. Those that achieved full marks were able to describe how this term refers to the usefulness of the study in the everyday lives of people.
- (b) Some candidates did well on this question and were able to think of a general area of application for each of the studies. For example, use of the Rosenhan study to improve diagnosis of patients going into a mental hospital. Very few candidates could give a sufficiently clear application for any of the studies. Many outlined the findings from the study without a comment on an application.

Candidates found the Freud study particularly difficult to apply to everyday life but most gained some credit for outlining detail from the study. Candidates often failed to mention the application to phobias in children.

- (c) For this question, candidates need to identify and discuss three problems with clear reference to a core study for each point. Many were able to describe one or two problems and included points related to ecological validity, generalisability and ethics. Unfortunately, many did not link their responses to a study and achieved fewer marks. Those who did provide links often did this in a very brief manner and achieved limited credit per problem raised.

# PSYCHOLOGY

---

Paper 9698/23

Core Studies 2

## Key Messages

### **Section A**

#### **Question 1**

Candidates should be aware of the requirements of each question in the exam. For example, if asked to describe they should not include evaluative comments. A few found it difficult to describe the snapshot method and how it was used in the Maguire et al. study in **part (a)** rather than evaluate the snapshot method. Candidates should suggest a simple alternative to the original study in **part (b)** and give clear details of the procedure followed including reference to the sampling method used. Extended evaluative points that make direct reference to the alternative idea are necessary in **part (c)** to achieve full marks.

#### **Question 2**

It is important that candidates practice writing these types of questions. Many did not structure their responses appropriately and could not achieve full marks. For example, if the question asks for strengths and weaknesses, at least two strengths and two weaknesses must be given with examples from the Loftus and Pickrell study.

Candidates must refer to the named study in their responses to all parts of the question to achieve higher marks. Some referred to studies other than Loftus and Pickrell and received no marks for these examples.

### **Section B**

Candidates must write more extended responses in both **part (b)** and **part (c)** of the essay as many gave accurate responses that lacked depth. Evidence must be given in **part (c)** to achieve higher marks.

## General comments

The marks achieved by candidates sitting this examination covered the whole spectrum of the mark scheme. Many provided good answers which showed that they were very well prepared and consistently referred to the evidence in order to achieve high marks.

Time management for this paper was good for most candidates and most attempted all questions that were required.

A minority of candidates answered both questions in the **Section B** essay. When a candidate did this they were awarded the mark for the better of the two questions (**Question 3** or **Question 4**). These candidates usually achieved poorly.

Candidates need to cover the entire syllabus so that they can respond to the questions in **Section A** where there is no choice of question. In addition to this, candidates must include evidence in the **part (c)** of their **Section B** essays to achieve higher marks. **Question 3** and **Question 4** were equally popular with candidates.

## **Comments on Specific Questions**

### **Section A**

#### **Question 1**

- (a) Many candidates did reasonably well on this question. Most were able to state that a snapshot study is brief and some gave an example. A few thought snapshot studies could go on for weeks or even months which is not correct. Many also gave clear examples from the Maguire et al. study and sometimes linked this to the snapshot method. A significant number evaluated in this part of the question and received no additional marks for this work.
- (b) Most candidates were able to describe a procedure using the longitudinal method and with participants other than taxi drivers. Popular ideas included using candidates about to take their driving test with a follow up later on to see if the hippocampus had altered in any way due to learning to drive or getting the participants to learn a particular route and describe this while in a scanner. Most candidates chose to use a PET or an MRI scanner. An MRI scanner was allowed however, if an EEG was suggested this was taken as a minor omission. A few failed to mention the scanner or did not carry out a longitudinal method and received limited credit for their procedure.

It was rare to find a fully replicable procedure. Candidates did not tend to refer to the sampling method and often did not name the scanning equipment used. Some did not give enough details of the tasks the participants were asked to do while in the scanner.

A very small minority of candidates evaluated their idea in this question and received no credit for this as this is the correct response to **Question 1(c)**.

- (c) The vast majority of candidates achieved marks in this question by providing some evaluative points.

Many discussed issues about the cost of using brain scan equipment, the time-consuming nature of the longitudinal method as well as the problems this might create with participant attrition. Many also brought in issues of ecological validity and generalisability of the limited sample used in their procedure.

Some candidates gave well developed points that achieved very high marks as they referred directly to their alternative idea. Some candidates only briefly identified issues and needed to refer back to the context of their own study.

#### **Question 2**

- (a) The majority of candidates achieved some credit for their response to this question. Most gave a very brief definition of reliability. Some achieved full marks by giving a more detailed definition or stating that reliability is about the consistency of the measuring device used.
- (b) Most candidates achieved some credit for their responses. Candidates were able to give a conclusion from the Loftus and Pickrell study and give a specific use of this conclusion in everyday life. Most referred to how it would be useful to know about the formation of false memories in the courtroom. Few were able to link the conclusion from the study to its use in everyday life.
- (c) The vast majority of candidates achieved some marks in this section. Most were able to describe at least one strength and one weakness of reliable research. Many referred to the study being easy to replicate in the future and also that there were good controls. For the weaknesses some referred to the lack of ecological validity or the frequent use of quantitative data which lack depth.

A number of candidates just evaluated the Loftus and Pickrell study and only strengths and weaknesses that were appropriate to reliable research were credited.

Some candidates attempted the ‘plural nature’ of these types of questions, describing two strengths and two weaknesses to achieve the higher marks.

- (d) A number of candidates were able to discuss the extent to which the Loftus and Pickrell study is useful. They referred to issues such as the lack of ecological validity, generalisability and the possible presence of demand characteristics that could have a negative impact on the useful of this piece of research. Most referred to the study in their response. However, many candidates did not answer the question set and described how the findings were useful without referring to the extent to which they are useful, and they gained limited credit.

## Section B

### Question 3

- (a) Most candidates achieved some credit for this question. They were able to state that generalisability refers to the ability to apply the results to the general population. Some referred to ecological validity and this was also creditworthy. A few candidates gave more detailed responses and described how generalisability refers to the extent to which the results of the sample used in the study can be applied to the target population.
- (b) This question was answered well by many candidates. Most achieved credit per study and many achieved full marks. Candidates were able to describe how the data were collected in each of the named studies. Many failed to describe the data collection fully and omitted important details from each description. For example, candidates needed to describe the duration and direction of gaze in the Langlois et al. study, exact details of the observations made in the Bandura et al. study and the different types of tests given to the kittens in the Held and Hein study.
- (c) Most candidates could identify many problems with making generalisations from psychological research. Popular points included ecological validity, generalising from the sample used in the study to the wider population and issues of ethics. Some candidates referred back to the core studies as evidence to back up their points. However, many did not do this and gained limited credit. Some candidates just made one point for all three studies. This was often done when referring to the problem of generalising from the sample. Candidates need to make three different points to achieve the maximum marks.

### Question 4

- (a) Many who attempted this question achieved some credit for stating that the situational explanation of behaviour refers to the effect the situation or environment has on a person's behaviour. Some achieved full marks, including an example of how the situation might affect behaviour.
- (b) Most candidates could achieve some credit by mentioning something correct about the study. Many attempted to describe how the situation affected the behaviour of the participants in the study but for higher marks most needed to make the links between the features of the situation (such as the drunk/cane victim in the Piliavin et al. study) and the behaviour shown by the participants. A few did this and achieved high marks.
- (c) Candidates responded well to this question and many described two or three problems that psychologists face when they investigate situational explanations of behaviour. Popular responses included ethical issues, problems with creating an ecologically valid situation and difficulties with generalisability of the samples used in the studies. Some needed to link their responses to a study to achieve higher marks.

# PSYCHOLOGY

---

Paper 9698/31  
Specialist Choices

## Key Messages

- Candidates should provide answers that equate to the mark allocation, so an answer worth 2 marks should be short and an answer worth 8 marks should be correspondingly longer.
- Candidates should appreciate that this is a three-hour examination and so it is expected that the amount of writing should be lengthy. A **Section B** essay (**parts (a) and (b)**) should take approximately 45–50 minutes and be at least 4 sides of paper in length.
- Candidates should read all parts of a question before beginning to answer to ensure that all parts of the question can be answered.
- Candidates should ensure that they know the difference between describe and evaluate for **Section B** questions and between describe and suggest for **Section C** questions.
- Candidates should look to quote psychological knowledge wherever possible. Anecdotal answers will not achieve top marks.
- Candidates should apply the methodological knowledge learned for Papers 1 and 2 not only to studies learned for Paper 3, but also to form the basis of their **Section C** suggestions.
- Candidates should always seek to evaluate using psychological methods, approaches, issues and debates as appear in the syllabus rather than with general evaluation points.

## General Comments

### **Section A (all options)**

A number of modifications to examination technique could improve marks:

- Writing an amount appropriate to the marks allocated. If a description of two studies is needed for 4 marks, the allocation of marks is 2 + 2, whereas if a description of one study is required for 4 marks, then the same amount in total should be written as for the 2 + 2 format.
- Writing an amount equivalent to 4 marks rather than 8 or 12 marks. Although there were many answers that were far too short, there were also many answers that were just as long as **Section B** essays.

### **Section B (all options)**

- Many candidates score full marks because they cover a range of bullet points on the syllabus, use psychological terminology, show accuracy in explanations, show they understand what they write and write an appropriate amount of detail. Candidates who partially address these aspects score mid-band marks and candidates who address one or two of these aspects score lower marks.
- An important message, which is always repeated, is that many answers would receive significantly higher marks if the difference between ‘describe’ and ‘evaluate’ is known by candidates. **Section B part (a)** is ‘describe’ and question **part (b)** is ‘evaluate’. Evaluation is not simply an additional description question part. Evaluation is a comment about what is good and what is not so good about the evidence that has been described in **part (a)**. Evaluation requires a candidate to think and apply and not to just reproduce learning.

Evaluation by candidates can often be divided into three types:

- those who evaluate using a number of evaluation issues in addition to the named issue (these candidates score the highest marks)
- those who focus exclusively on the one named issue and gain limited marks, because one issue is not a range (as required by the mark scheme), and those who exclude the named issue altogether (also gain limited marks)
- those who do not evaluate at all (and score no marks) merely describing more information.

For many candidates of the first type, answers could be improved further by taking time to clarify what they write. For example, a candidate might write ‘disruptive behaviour is reliable’ and ‘disruptive behaviour is valid’, both of which have very little meaning. A little elaboration would show understanding and that a candidate has taken time to think about what is written. This is preferable than candidates who reproduce rote learning without demonstrating understanding.

### **Section C (all options)**

One question part asks a candidate to describe and the one question part asks a candidate to suggest. Many candidates appear not to know the difference between these two. Description is to show knowledge and understanding that has been learned. A suggestion is to go beyond description and to think about how something could be investigated (studied) or applied to a given situation. A suggestion is not something that can be learned beforehand. It is requiring a candidate to think for themselves during the examination.

When a question asks candidates to use a specific method, then that method must be used. Some candidates start with ‘I will conduct an experiment’ and write nothing further about the IV or DV or controls or apply a design (repeated measures, for example). These are essential features of an experiment and should be included.

Candidates should show their methodological knowledge because many marks can be gained for application of this knowledge in this section.

### **COMMENTS ON SPECIFIC QUESTIONS**

#### **PSYCHOLOGY AND EDUCATION**

##### **Question 1**

- (a) Nearly all candidates scored credit when they provided a basic answer, saying for example that extrinsic motivation is due to an external reward. Many candidates provided some elaboration such as an example, the most common being praise from a teacher.
- (b) The most commonly described theory was that of behaviourism where positive reinforcement increases the probability of a behaviour happening again. Many candidates referred to the work of Brophy who looks at the use of effective and ineffective praise. Candidates could be distinguished by the amount of detail and accuracy in their answers.

##### **Question 2**

- (a) Some candidates produced superb answers which were worth maximum marks. What characterised these answers was the range of information, the use of relevant psychological examples, the detail included and the depth of understanding. Other candidates provided some detail and some studies but were often lacking in a number of important aspects needed for top band marks (see mark scheme for details).
- (b) Some candidates provided superb evaluations showing abundant evidence of all the necessary skills. However, some candidates did not evaluate, but described more of what should be included in (a). To achieve maximum marks, candidates need to include a range of evaluation issues, one of which should be the named issue; in this case quantitative data. Some candidates did not mention this issue at all, although IQ tests use quantitative data.

##### **Question 3**

- (a) Some candidates gave anecdotal answers that didn’t answer the question set, such as comment that a good teacher is needed, or that all that students should concentrate hard and take good notes in class, which could not be credited. Candidates scoring maximum marks often described McCarthy’s 4-Mat system, Mulcahy’s SPELT, or PQRST. As three marks were allocated to each description, candidates did not need very much detail on each, and some candidates wrote far too much. Candidates are reminded to think about how much they should write in relation to the allocation of marks (see key messages).

- (b) The most logical method to use was a field experiment and this was used by most candidates. Some candidates suggested using several different methods, but this is not advisable because the details of any one method are often written superficially. Candidates scoring top marks demonstrated knowledge of IV, DV, controls and designs and drew a conclusion about the most effective study skill from the data their study would have gathered.

**Question 4**

- (a) Some candidates appeared to be unfamiliar with observations. Other candidates identified a type of observation, with some choosing participant observation and some choosing non-participant natural observation. These candidates often identified response categories (for example different types of bullying and the frequency of each) and considered using two observers so the reliability of the data gathered could be checked.
- (b) Most candidates could give two explanations for bullying but sometimes these consisted of a single sentence only. Other candidates provided much more detail and often quoted appropriate psychological research to support their answer.

**PSYCHOLOGY AND HEALTH**

**Question 5**

- (a) Many candidates could not provide an adequate answer, often guessing and often being incorrect. Some candidates made a distinction between 'dayshift' and 'nightshift' and this was credited. One definition of shiftwork is that it is the division of a 24-hour work schedule into smaller workable time periods such as eight hours, each time period being called a shift.
- (b) Those candidates struggling to answer part (a) also struggled with this question part. For those candidates who had studied shiftwork, maximum marks were often scored when they described slow rotation theory and then suggested rapid rotation theory as the alternative way to organise shiftwork.

**Question 6**

- (a) Answers to this question covered the entire mark range. Some candidates included no relevant psychology and wrote anecdotal answers. Other candidates focused on health as that specifically experienced in a practitioner-patient relationship rather than what was recommended for study on the syllabus. However, there were superb answers which were well written, showed a competent knowledge of published research studies, and appropriate examination technique was evident.
- (b) The named evaluation issue on which candidates had to focus for one of their evaluation issues was snapshot and longitudinal studies. This should have provided good discussion given that some studies (on fear arousal for example) were snapshot compared to others which were longitudinal (such as those on health in schools, communities, and those by Lewin et al. on heart health). Some candidates included only this issue and gained limited credit (see mark scheme an explanation), whilst others included three or more issues and often scored full marks.

**Question 7**

- (a) Candidates were required to use a questionnaire to investigate the customising of treatment in non-adherence, but many candidates suggested methods other than a questionnaire, which could not be credited. Some candidates suggested questionnaires but demonstrated little knowledge about questionnaire design. At the top end of the mark range candidates often suggested starting with closed questions and gathering quantitative data and ending with open ended questions to gather additional qualitative data. The best candidates also gave examples of questions that could be asked and specified the type of rating scale for the participants to respond.
- (b) The syllabus gives the Johnson and Bytheway study as an example for the topic of customising treatment. If this study is not taught then an appropriate alternative must replace it. It is insufficient to have one or two anecdotal answers replacing published research. This meant that some candidates scored very few marks. Some candidates described the Johnson and Bytheway research in detail and some were able to gain full marks.

### Question 8

- (a) Phantom limb pain occurs when a limb has been amputated but a person still experiences pain where the amputated limb once was. This appeared to confuse a number of candidates because they suggested using a drawing and colouring in the place on the body is where the pain is. Candidates at the top end of the mark range suggested using a clinical interview or suggested using parts of an MPQ (McGill pain questionnaire) such as the part asking the person to describe the sensation of the pain. Other candidates suggested using a type of visual analogue scale.
- (b) Some candidates did not know what psychogenic pain is and could not example of it, despite phantom limb pain being specified in **part (a)**. Other candidates adequately described that psychogenic pain is experienced through something other than a physical injury and that phantom limb pain is a perfect example.

## **PSYCHOLOGY AND ENVIRONMENT**

### Question 9

- (a) To score full marks candidates had to address both the terms ‘crowding’ and ‘human health’. Many candidates failed to explain the term crowding and also sometimes ignored the question and wrote about the effects on pro-social behaviour or of performance, and scored no marks. The correct answer was the crowding is the perception of restrictions of physical space, and one effect of this is that it can have a negative effect on health, such as causing stress.
- (b) Most candidates chose to describe the study by Lundberg (1976) which was conducted on a commuter train in Sweden, although the study by Evans and Wener was also mentioned. The results of both these studies showed that crowding had a negative effect on health. Marks were allocated according to the accuracy and detail of description along with depth of understanding.

### Question 10

- (a) Candidates answering this question were of two types: those who could not demonstrate knowledge about architecture and often wrote answers scoring no more than one or two marks; those candidates who demonstrated significant knowledge about architecture and described theories of urban living, urban renewal and housing design, and also community environmental design. This latter category of candidates often scored maximum marks.
- (b) Those candidates scoring maximum marks in (a) often scored high marks in (b). These candidates included a range of issues, one of which being the named issue of individual and situational explanations. These answers often included appropriate advantages and disadvantages which were supported with relevant examples. Those candidates scoring low marks in (a) sometimes scored a few marks, but struggled when they were unable to apply any relevant example to their answer.

### Question 11

- (a) A free choice of method was given in this question. Some candidates chose to give a questionnaire to local residents whilst other candidates suggested interviewing local residents. Some candidates suggested conducting laboratory experiments to determine levels of acceptable and unacceptable noise. Other candidates suggested conducting a field experiment and gathering data on actual aircraft noise levels. What distinguished top marks from those in lower bands was the extent of methodological knowledge evident in answers.
- (b) The study by Bronzaft was most commonly quoted with marks being allocated according to the accuracy of the detail about the study. Also commonly quoted were the studies by Haines et al. on aircraft noise which also looked at the performance of children in schools. Some candidates quoted the study by Glass and Singer on transportation noise of cars but this study did not focus specifically on performance.

### Question 12

- (a) This question required candidates to specifically conduct an interview. Some candidates wrote excellent answers and showed an in-depth knowledge of types of interview, types of question, and how data could be gathered. Other candidates did not demonstrate knowledge of interviews as a method. If nothing is known about the named method then the alternative question appearing in this section should be considered.
- (b) The London bombing appears on the syllabus under the bullet heading of 'psychological intervention before and after events', the Rubin et al. study is the required study for this topic. As was mentioned in part (a) above, if nothing is known about the London bombings (or any other aspect asked about in a **Section C** question, then the alternative question in this section should be attempted. Many candidates were very familiar with the London bombings of 2005 and the Rubin et al. study and wrote excellent answers.

## PSYCHOLOGY AND ABNORMALITY

### Question 13

- (a) To score full marks for this question part candidates had to include in their answer a mention of both 'case study' and 'phobia'. Most candidates scored some credit but rarely the maximum when failing to address both components. A case study is usually an in-depth study of a single person (in abnormal psychology), and a phobia is an irrational fear of something.
- (b) Nearly all candidates scored full marks, often describing in sufficient detail the study by Watson on little Albert, by Freud on little Hans, and the study by McGrath et al. of Lucy who had a fear of loud bangs such as popping balloons and fireworks.

### Question 14

- (a) Answers to this question could include definitions of abnormality such as deviation from statistical norms etc., and models of abnormality themselves such as the medical, behavioural, psychodynamic and cognitive models. Candidates could also include treatments of abnormality based on models. Whereas many candidates included all of these aspects in their answers, and showed very good understanding of different models, other candidates focused only on definitions and this gained limited marks.
- (b) For this question the named issue was the usefulness of models. This should have been just one of at least three issues included in answers. However, some candidates only included the named issue and even though there were some brilliant answers were written, these gained limited marks (see mark scheme). Many other candidates included least three issues, provided detailed discussion and were awarded high marks.

### Question 15

- (a) A number of candidates decided to replicate a version of the Rosenhan study to see how many psychiatrists would diagnose them as having schizophrenia. It was often unclear how this would test knowledge of different types or cover a sufficiently large sample. Other candidates suggested conducting a questionnaire that could be sent to large numbers of psychiatrists. Some of these answers gave no details about the questions to be asked whereas other candidates asked a simple question such as 'do you know that types of schizophrenia are no longer on DSM?', to be answered yes or no. Candidates showing appropriate methodological knowledge scored higher marks.
- (b) Some candidates gave details of the different types of schizophrenia while others did not. Some candidates outlined positive and negative symptoms while others did not. Some candidates described the more general features of schizophrenia, such as sufferers having an impaired sense of reality, whilst other candidates mentioned more specific features.

### Question 16

- (a) In order to score marks candidates needed to know about both cognitive behaviour therapy (CBT) and some details about the disorder they would treat. Some candidates knew very little about CBT and scored low marks. Other candidates had extensive knowledge and made very good and entirely appropriate suggestions.
- (b) This question part wanted a description of the theory on which cognitive behaviour therapy is based. Nearly all candidates gave appropriate descriptions with only one or two candidates confusing it with psychoanalytic therapy. A few candidates wrote about the cognitive therapies of Ellis and Beck but this isn't CBT. Some candidates thought that CBT involved giving rewards for appropriate behaviour and a few even suggested giving rewards for not engaging in the behaviour that was being treated.

## PSYCHOLOGY AND ORGANISATIONS

### Question 17

- (a) Many candidates failed to provide an adequate explanation of the term 'ergonomics', but the term is the title of a bullet point of the syllabus and candidates should be familiar with it. A full definition of ergonomics is 'the applied science of equipment designed to maximise comfort and ease of use for the employee and to maximise efficiency productivity for the employer'. Any attempt at this definition would score marks.
- (b) A few candidates scored maximum marks, other candidates did not know the term ergonomics or what was meant by operator-machine systems, and so these answers scored no marks. It is essential that candidates' preparation covers the entire syllabus.

### Question 18

- (a) While there were some very high quality answers, there were also many weak answers. Some candidates focused only on decision making and autocratic and democratic decisions. Other candidates focused only on leadership style and the work of Fiedler. Answers need to cover a range of information to access top band marks. Answers at the top end of the mark range included a number of different theories of leadership, considered leadership style and effectiveness, and also considered the relationship between leaders and followers.
- (b) Marks here covered the entire mark range, distinguishing between those candidates evaluated successfully and provided detailed competent answers, and those who did not evaluate. The named issue, which must be addressed, was a discussion about nature versus nurture. This should have been straightforward given some theories claim that leaders are born and other theories claim that leaders are made. However, a small number of candidates did not discuss this issue and some appeared not to know the terms nature and nurture.

### Question 19

- (a) The most logical design would be to have different strategies as conditions of an independent variable, for some task to be set and then the dependent variable being some measure on which the strategies could be judged. An answer like this would show knowledge of methodology and of goal-setting strategies. Whereas some candidates did this successfully many others did not. A few candidates suggested different methods such as observations and questionnaires which could not be credited. If the question states 'experiment' then the suggestion must be an experiment.
- (b) Some candidates could not demonstrate any knowledge of the goal-setting theory proposed by Latham and Locke, others managed a few sentences, whilst others provided excellent detail that was accurate and showed understanding and scored full marks. The theory has five main principles which are: clarity, challenge, commitment, that the goals be effective, and that the goals be achievable.

**Question 20**

- (a) This question invited candidates to assess the change in psychological state, such as mood or happiness, of workers who had undergone a job rotation. A free choice of method was given. Many candidates suggested that an interview or questionnaire be conducted with workers. Marks were allocated according to the methodological knowledge demonstrated. For example the type of interview, the type of questionnaire, examples of questions to be asked, and how answers in response to the questions could be recorded should have been included to access top band marks.
- (b) While the job characteristics model proposed by Hackman and Oldham is suggested on the syllabus candidates are free to describe any other job characteristics model. The Hackman and Oldham model looks at both the characteristics of a job and the psychological state of workers, dividing both of these into a number of categories. While many candidates knew this model, others guessed and provided entirely anecdotal answers.

# PSYCHOLOGY

---

Paper 9698/32  
Specialist Choices

## Key Messages

- Candidates should provide answers that equate to the mark allocation, so an answer worth 2 marks should be short and an answer worth 8 marks should be correspondingly longer.
- Candidates should appreciate that this is a three-hour examination and so it is expected that the amount of writing should be lengthy. A **Section B** essay (**parts (a) and (b)**) should take approximately 45–50 minutes and be at least 4 sides of paper in length.
- Candidates should read all parts of a question before beginning to answer to ensure that all parts of the question can be answered.
- Candidates should ensure that they know the difference between describe and evaluate for **Section B** questions and between describe and suggest for **Section C** questions.
- Candidates should look to quote psychological knowledge wherever possible. Anecdotal answers will not achieve top marks.
- Candidates should apply the methodological knowledge learned for Papers 1 and 2 not only to studies learned for Paper 3, but also to form the basis of their **Section C** suggestions.
- Candidates should always seek to evaluate using psychological methods, approaches, issues and debates as appear in the syllabus rather than with general evaluation points.

## General Comments

### **Section A (all options)**

A number of modifications to examination technique could improve marks:

- Writing an amount appropriate to the marks allocated. If a description of two studies is needed for 4 marks, the allocation of marks is 2 + 2, whereas if a description of one study is required for 4 marks, then the same amount in total should be written as for the 2 + 2 format.
- Writing an amount equivalent to 4 marks rather than 8 or 12 marks. Although there were many answers that were far too short, there were also many answers that were just as long as **Section B** essays.

### **Section B (all options)**

- Many candidates score full marks because they cover a range of bullet points on the syllabus, use psychological terminology, show accuracy in explanations, show they understand what they write and write an appropriate amount of detail. Candidates who partially address these aspects score mid-band marks and candidates who address one or two of these aspects score lower marks.
- An important message, which is always repeated, is that many answers would receive significantly higher marks if the difference between ‘describe’ and ‘evaluate’ is known by candidates. **Section B part (a)** is ‘describe’ and question **part (b)** is ‘evaluate’. Evaluation is not simply an additional description question part. Evaluation is a comment about what is good and what is not so good about the evidence that has been described in **part (a)**. Evaluation requires a candidate to think and apply and not to just reproduce learning.

Evaluation by candidates can often be divided into three types:

- those who evaluate using a number of evaluation issues in addition to the named issue (these candidates score the highest marks)
- those who focus exclusively on the one named issue and gain limited marks, because one issue is not a range (as required by the mark scheme), and those who exclude the named issue altogether (also gain limited marks)
- those who do not evaluate at all (and score no marks) merely describing more information.

For many candidates of the first type, answers could be improved further by taking time to clarify what they write. For example, a candidate might write ‘disruptive behaviour is reliable’ and ‘disruptive behaviour is valid’, both of which have very little meaning. A little elaboration would show understanding and that a candidate has taken time to think about what is written. This is preferable than candidates who reproduce rote learning without demonstrating understanding.

### **Section C (all options)**

One question part asks a candidate to describe and the one question part asks a candidate to suggest. Many candidates appear not to know the difference between these two. Description is to show knowledge and understanding that has been learned. A suggestion is to go beyond description and to think about how something could be investigated (studied) or applied to a given situation. A suggestion is not something that can be learned beforehand. It is requiring a candidate to think for themselves during the examination.

When a question asks candidates to use a specific method, then that method must be used. Some candidates start with ‘I will conduct an experiment’ and write nothing further about the IV or DV or controls or apply a design (repeated measures, for example). These are essential features of an experiment and should be included.

Candidates should show their methodological knowledge because many marks can be gained for application of this knowledge in this section.

### **COMMENTS ON SPECIFIC QUESTIONS**

#### **PSYCHOLOGY AND EDUCATION**

##### **Question 1**

- (a) Nearly all candidates scored credit when they provided a basic answer, saying for example that extrinsic motivation is due to an external reward. Many candidates provided some elaboration such as an example, the most common being praise from a teacher.
- (b) The most commonly described theory was that of behaviourism where positive reinforcement increases the probability of a behaviour happening again. Many candidates referred to the work of Brophy who looks at the use of effective and ineffective praise. Candidates could be distinguished by the amount of detail and accuracy in their answers.

##### **Question 2**

- (a) Some candidates produced superb answers which were worth maximum marks. What characterised these answers was the range of information, the use of relevant psychological examples, the detail included and the depth of understanding. Other candidates provided some detail and some studies but were often lacking in a number of important aspects needed for top band marks (see mark scheme for details).
- (b) Some candidates provided superb evaluations showing abundant evidence of all the necessary skills. However, some candidates did not evaluate, but described more of what should be included in (a). To achieve maximum marks, candidates need to include a range of evaluation issues, one of which should be the named issue; in this case quantitative data. Some candidates did not mention this issue at all, although IQ tests use quantitative data.

##### **Question 3**

- (a) Some candidates gave anecdotal answers that didn’t answer the question set, such as comment that a good teacher is needed, or that all that students should concentrate hard and take good notes in class, which could not be credited. Candidates scoring maximum marks often described McCarthy’s 4-Mat system, Mulcahy’s SPELT, or PQRST. As three marks were allocated to each description, candidates did not need very much detail on each, and some candidates wrote far too much. Candidates are reminded to think about how much they should write in relation to the allocation of marks (see key messages).

- (b) The most logical method to use was a field experiment and this was used by most candidates. Some candidates suggested using several different methods, but this is not advisable because the details of any one method are often written superficially. Candidates scoring top marks demonstrated knowledge of IV, DV, controls and designs and drew a conclusion about the most effective study skill from the data their study would have gathered.

**Question 4**

- (a) Some candidates appeared to be unfamiliar with observations. Other candidates identified a type of observation, with some choosing participant observation and some choosing non-participant natural observation. These candidates often identified response categories (for example different types of bullying and the frequency of each) and considered using two observers so the reliability of the data gathered could be checked.
- (b) Most candidates could give two explanations for bullying but sometimes these consisted of a single sentence only. Other candidates provided much more detail and often quoted appropriate psychological research to support their answer.

**PSYCHOLOGY AND HEALTH**

**Question 5**

- (a) Many candidates could not provide an adequate answer, often guessing and often being incorrect. Some candidates made a distinction between 'dayshift' and 'nightshift' and this was credited. One definition of shiftwork is that it is the division of a 24-hour work schedule into smaller workable time periods such as eight hours, each time period being called a shift.
- (b) Those candidates struggling to answer part (a) also struggled with this question part. For those candidates who had studied shiftwork, maximum marks were often scored when they described slow rotation theory and then suggested rapid rotation theory as the alternative way to organise shiftwork.

**Question 6**

- (a) Answers to this question covered the entire mark range. Some candidates included no relevant psychology and wrote anecdotal answers. Other candidates focused on health as that specifically experienced in a practitioner-patient relationship rather than what was recommended for study on the syllabus. However, there were superb answers which were well written, showed a competent knowledge of published research studies, and appropriate examination technique was evident.
- (b) The named evaluation issue on which candidates had to focus for one of their evaluation issues was snapshot and longitudinal studies. This should have provided good discussion given that some studies (on fear arousal for example) were snapshot compared to others which were longitudinal (such as those on health in schools, communities, and those by Lewin et al. on heart health). Some candidates included only this issue and gained limited credit (see mark scheme an explanation), whilst others included three or more issues and often scored full marks.

**Question 7**

- (a) Candidates were required to use a questionnaire to investigate the customising of treatment in non-adherence, but many candidates suggested methods other than a questionnaire, which could not be credited. Some candidates suggested questionnaires but demonstrated little knowledge about questionnaire design. At the top end of the mark range candidates often suggested starting with closed questions and gathering quantitative data and ending with open ended questions to gather additional qualitative data. The best candidates also gave examples of questions that could be asked and specified the type of rating scale for the participants to respond.
- (b) The syllabus gives the Johnson and Bytheway study as an example for the topic of customising treatment. If this study is not taught then an appropriate alternative must replace it. It is insufficient to have one or two anecdotal answers replacing published research. This meant that some candidates scored very few marks. Some candidates described the Johnson and Bytheway research in detail and some were able to gain full marks.

### Question 8

- (a) Phantom limb pain occurs when a limb has been amputated but a person still experiences pain where the amputated limb once was. This appeared to confuse a number of candidates because they suggested using a drawing and colouring in the place on the body is where the pain is. Candidates at the top end of the mark range suggested using a clinical interview or suggested using parts of an MPQ (McGill pain questionnaire) such as the part asking the person to describe the sensation of the pain. Other candidates suggested using a type of visual analogue scale.
- (b) Some candidates did not know what psychogenic pain is and could not example of it, despite phantom limb pain being specified in **part (a)**. Other candidates adequately described that psychogenic pain is experienced through something other than a physical injury and that phantom limb pain is a perfect example.

## **PSYCHOLOGY AND ENVIRONMENT**

### Question 9

- (a) To score full marks candidates had to address both the terms ‘crowding’ and ‘human health’. Many candidates failed to explain the term crowding and also sometimes ignored the question and wrote about the effects on pro-social behaviour or of performance, and scored no marks. The correct answer was the crowding is the perception of restrictions of physical space, and one effect of this is that it can have a negative effect on health, such as causing stress.
- (b) Most candidates chose to describe the study by Lundberg (1976) which was conducted on a commuter train in Sweden, although the study by Evans and Wener was also mentioned. The results of both these studies showed that crowding had a negative effect on health. Marks were allocated according to the accuracy and detail of description along with depth of understanding.

### Question 10

- (a) Candidates answering this question were of two types: those who could not demonstrate knowledge about architecture and often wrote answers scoring no more than one or two marks; those candidates who demonstrated significant knowledge about architecture and described theories of urban living, urban renewal and housing design, and also community environmental design. This latter category of candidates often scored maximum marks.
- (b) Those candidates scoring maximum marks in (a) often scored high marks in (b). These candidates included a range of issues, one of which being the named issue of individual and situational explanations. These answers often included appropriate advantages and disadvantages which were supported with relevant examples. Those candidates scoring low marks in (a) sometimes scored a few marks, but struggled when they were unable to apply any relevant example to their answer.

### Question 11

- (a) A free choice of method was given in this question. Some candidates chose to give a questionnaire to local residents whilst other candidates suggested interviewing local residents. Some candidates suggested conducting laboratory experiments to determine levels of acceptable and unacceptable noise. Other candidates suggested conducting a field experiment and gathering data on actual aircraft noise levels. What distinguished top marks from those in lower bands was the extent of methodological knowledge evident in answers.
- (b) The study by Bronzaft was most commonly quoted with marks being allocated according to the accuracy of the detail about the study. Also commonly quoted were the studies by Haines et al. on aircraft noise which also looked at the performance of children in schools. Some candidates quoted the study by Glass and Singer on transportation noise of cars but this study did not focus specifically on performance.

### Question 12

- (a) This question required candidates to specifically conduct an interview. Some candidates wrote excellent answers and showed an in-depth knowledge of types of interview, types of question, and how data could be gathered. Other candidates did not demonstrate knowledge of interviews as a method. If nothing is known about the named method then the alternative question appearing in this section should be considered.
- (b) The London bombing appears on the syllabus under the bullet heading of 'psychological intervention before and after events', the Rubin et al. study is the required study for this topic. As was mentioned in part (a) above, if nothing is known about the London bombings (or any other aspect asked about in a **Section C** question, then the alternative question in this section should be attempted. Many candidates were very familiar with the London bombings of 2005 and the Rubin et al. study and wrote excellent answers.

## PSYCHOLOGY AND ABNORMALITY

### Question 13

- (a) To score full marks for this question part candidates had to include in their answer a mention of both 'case study' and 'phobia'. Most candidates scored some credit but rarely the maximum when failing to address both components. A case study is usually an in-depth study of a single person (in abnormal psychology), and a phobia is an irrational fear of something.
- (b) Nearly all candidates scored full marks, often describing in sufficient detail the study by Watson on little Albert, by Freud on little Hans, and the study by McGrath et al. of Lucy who had a fear of loud bangs such as popping balloons and fireworks.

### Question 14

- (a) Answers to this question could include definitions of abnormality such as deviation from statistical norms etc., and models of abnormality themselves such as the medical, behavioural, psychodynamic and cognitive models. Candidates could also include treatments of abnormality based on models. Whereas many candidates included all of these aspects in their answers, and showed very good understanding of different models, other candidates focused only on definitions and this gained limited marks.
- (b) For this question the named issue was the usefulness of models. This should have been just one of at least three issues included in answers. However, some candidates only included the named issue and even though there were some brilliant answers were written, these gained limited marks (see mark scheme). Many other candidates included least three issues, provided detailed discussion and were awarded high marks.

### Question 15

- (a) A number of candidates decided to replicate a version of the Rosenhan study to see how many psychiatrists would diagnose them as having schizophrenia. It was often unclear how this would test knowledge of different types or cover a sufficiently large sample. Other candidates suggested conducting a questionnaire that could be sent to large numbers of psychiatrists. Some of these answers gave no details about the questions to be asked whereas other candidates asked a simple question such as 'do you know that types of schizophrenia are no longer on DSM?', to be answered yes or no. Candidates showing appropriate methodological knowledge scored higher marks.
- (b) Some candidates gave details of the different types of schizophrenia while others did not. Some candidates outlined positive and negative symptoms while others did not. Some candidates described the more general features of schizophrenia, such as sufferers having an impaired sense of reality, whilst other candidates mentioned more specific features.

### Question 16

- (a) In order to score marks candidates needed to know about both cognitive behaviour therapy (CBT) and some details about the disorder they would treat. Some candidates knew very little about CBT and scored low marks. Other candidates had extensive knowledge and made very good and entirely appropriate suggestions.
- (b) This question part wanted a description of the theory on which cognitive behaviour therapy is based. Nearly all candidates gave appropriate descriptions with only one or two candidates confusing it with psychoanalytic therapy. A few candidates wrote about the cognitive therapies of Ellis and Beck but this isn't CBT. Some candidates thought that CBT involved giving rewards for appropriate behaviour and a few even suggested giving rewards for not engaging in the behaviour that was being treated.

## PSYCHOLOGY AND ORGANISATIONS

### Question 17

- (a) Many candidates failed to provide an adequate explanation of the term 'ergonomics', but the term is the title of a bullet point of the syllabus and candidates should be familiar with it. A full definition of ergonomics is 'the applied science of equipment designed to maximise comfort and ease of use for the employee and to maximise efficiency productivity for the employer'. Any attempt at this definition would score marks.
- (b) A few candidates scored maximum marks, other candidates did not know the term ergonomics or what was meant by operator-machine systems, and so these answers scored no marks. It is essential that candidates' preparation covers the entire syllabus.

### Question 18

- (a) While there were some very high quality answers, there were also many weak answers. Some candidates focused only on decision making and autocratic and democratic decisions. Other candidates focused only on leadership style and the work of Fiedler. Answers need to cover a range of information to access top band marks. Answers at the top end of the mark range included a number of different theories of leadership, considered leadership style and effectiveness, and also considered the relationship between leaders and followers.
- (b) Marks here covered the entire mark range, distinguishing between those candidates evaluated successfully and provided detailed competent answers, and those who did not evaluate. The named issue, which must be addressed, was a discussion about nature versus nurture. This should have been straightforward given some theories claim that leaders are born and other theories claim that leaders are made. However, a small number of candidates did not discuss this issue and some appeared not to know the terms nature and nurture.

### Question 19

- (a) The most logical design would be to have different strategies as conditions of an independent variable, for some task to be set and then the dependent variable being some measure on which the strategies could be judged. An answer like this would show knowledge of methodology and of goal-setting strategies. Whereas some candidates did this successfully many others did not. A few candidates suggested different methods such as observations and questionnaires which could not be credited. If the question states 'experiment' then the suggestion must be an experiment.
- (b) Some candidates could not demonstrate any knowledge of the goal-setting theory proposed by Latham and Locke, others managed a few sentences, whilst others provided excellent detail that was accurate and showed understanding and scored full marks. The theory has five main principles which are: clarity, challenge, commitment, that the goals be effective, and that the goals be achievable.

**Question 20**

- (a) This question invited candidates to assess the change in psychological state, such as mood or happiness, of workers who had undergone a job rotation. A free choice of method was given. Many candidates suggested that an interview or questionnaire be conducted with workers. Marks were allocated according to the methodological knowledge demonstrated. For example the type of interview, the type of questionnaire, examples of questions to be asked, and how answers in response to the questions could be recorded should have been included to access top band marks.
- (b) While the job characteristics model proposed by Hackman and Oldham is suggested on the syllabus candidates are free to describe any other job characteristics model. The Hackman and Oldham model looks at both the characteristics of a job and the psychological state of workers, dividing both of these into a number of categories. While many candidates knew this model, others guessed and provided entirely anecdotal answers.

# PSYCHOLOGY

Paper 9698/33  
Specialist Choices

## Key Messages

- Candidates should provide answers that equate to the mark allocation, so an answer worth 2 marks should be short and an answer worth 8 marks should be correspondingly longer.
- Candidates should appreciate that this is a three hour examination and so it is expected that the amount of writing should be lengthy. A **Section B** essay (**parts (a) and (b)**) should take approximately 45–50 minutes and be at least 4 sides of paper in length.
- Candidates should read all parts of a question before beginning to answer to ensure that all parts of the question can be answered.
- Candidates should ensure that they know the difference between describe and evaluate for **Section B** questions and between describe and suggest for **Section C** questions.
- Candidates should look to quote psychological knowledge wherever possible. Anecdotal answers will not achieve top marks.
- Candidates should apply the methodological knowledge learned for Papers 1 and 2 not only to studies learned for Paper 3, but also to form the basis of their **Section C** suggestions.
- Candidates should always seek to evaluate using psychological methods, approaches, issues and debates as appear in the syllabus rather than with general evaluation points.

## General Comments

### **Section A (all options)**

A number of modifications to examination technique could improve marks:

- Writing an amount appropriate to the marks allocated. If a description of two studies is needed for 4 marks, the allocation of marks is 2 + 2, whereas if a description of one study is required for 4 marks, then the same amount in total should be written as for the 2 + 2 format.
- Writing an amount equivalent to 4 marks rather than 8 or 12 marks. Although there were many answers that were far too short, there were also many answers that were just as long as **Section B** essays.

### **Section B (all options)**

- Many candidates score full marks because they cover a range of bullet points on the syllabus, use psychological terminology, show accuracy in explanations, show they understand what they write and write an appropriate amount of detail. Candidates who partially address these aspects score mid-band marks and candidates who address one or two of these aspects score just one or two marks.
- An important message, which is always repeated, is that many answers would receive significantly higher marks if the difference between ‘describe’ and ‘evaluate’ is known by candidates. **Section B question part (a)** will always be ‘describe’ and **question part (b)** will always be ‘evaluate’. Evaluation is not simply an additional description section as many candidates taking this paper appeared to think. Evaluation is a comment about what is good and what is not so good about the evidence that has been described in **part (a)**. Evaluation requires a candidate to think and apply and not to just reproduce learning.

Evaluation by candidates can often be divided into three types:

- those who evaluate using a number of evaluation issues in addition to the named issue (and these candidates score the highest marks)

- those who focus exclusively on the one named issue and have marks restricted, because one issue is not a range (as required by the mark scheme), and those who exclude the named issue altogether (and also have marks restricted)
- those who do not evaluate at all (and score no marks) merely describing more information.

For many candidates of the first type, answers could be improved further by taking a few moments to clarify what they write. For example, a candidate might write “disruptive behaviour is reliable” and “disruptive behaviour is valid” both of which have very little meaning. A little elaboration would show understanding and that a candidate has taken time to think about what is written. This is far preferable that candidates who blindly reproduce what they have learned in vast detail with little understanding of what any of it might mean.

### **Section C (all options)**

One question part asks a candidate to describe and the second question part asks a candidate to suggest. Many candidates appear not to know the difference between these two. Description is to show knowledge and understanding that has been learned. A suggestion is to go beyond description and to think about how something could be investigated (studied) or applied to a given situation. A suggestion is not something that can be learned beforehand. It is requiring a candidate to think for themselves during the examination.

When a question asks candidates to use a specific method, then that method must be used. Candidates often start with ‘I will conduct an experiment’ and write nothing further about the IV or DV or controls or apply a design (repeated measures, for example). These are essential features of an experiment and should be included.

Candidates should show their methodological knowledge because application of it is what scores most marks in this Section.

### **Comments on Specific Questions**

#### **PSYCHOLOGY AND EDUCATION**

Too few candidates to write a meaningful report.

#### **PSYCHOLOGY AND HEALTH**

##### **Question 5**

- (a) All candidates scored at least some credit, with many candidates scoring maximum marks. Those scoring limited marks usually wrote no more than a few words, such as “Munchausen syndrome is where people seek out excessive medical attention”, whereas those scoring full marks elaborated beyond the basic by giving an example or explaining the term.
- (b) Candidates were required to describe one study of Munchausen syndrome. The syllabus recommends the study by Aleem and Ajarim, but any other relevant study would score marks as this study is listed in the syllabus as an example. Many candidates described the Aleem and Ajarim case study in detail, often quoting precise detail. Some candidates wrote quite brief answers and as a result scored fewer marks. Some candidates described case of Beverly Allitt, a nurse who murdered four children. Although this is a case of Munchausen syndrome *by proxy* such answers still received full credit.

### Question 6

- (a) Some candidates wrote excellent answers that were very detailed and covered a range of different aspects of adherence listed on the syllabus and showed very clear understanding of what was written. For example, candidates structured their answers to include types of non-adherence, measures of non-adherence, and how adherence can be improved. In contrast some candidates wrote answers that were based on patient-practitioner relationships such as the work of McKinstry and Wang, and others wrote about doctor-centred and patient-centred relationships. While these studies do have an impact on patient adherence, and credit was awarded for such research, they offer nothing more than 'adherence is more or less likely' and do not cover the same range (such as measures) of research/studies as that appear on the syllabus for adherence.
- (b) Many candidates evaluated competently and thoroughly considering a range of appropriate issues. However a number of candidates provided no evaluation, but described what had already been written in (a) in more detail. Candidates should evaluate using a range of relevant issues, and include the named issue. For example, for this question the named issue required candidates to evaluate different methodologies used to study adherence. These methodologies could include the different ways adherence can be measured: through subjective self-reports, objective pill-counts, or pharmacy repeat prescriptions, in addition to others. Notably those candidates opting for the patient-practitioner approach could not score marks because those methodologies were not assessing adherence.

### Question 7

- (a) A number of candidates did not suggest an experiment (as the question specified), instead writing about observations or questionnaires. It is essential that candidates answer the question set, because if they do not then no marks can be awarded. Some candidates wrote about stress itself rather than focusing specifically on stress inoculation. For example candidates would look at measures of stress and causes of stress rather than the specific section of the syllabus on preventing stress. Those candidates who focused on stress inoculation specifically, and designed an experiment, wrote some innovative and interesting answers which often scored high marks.
- (b) Many candidates focused on the work of Meichenbaum who outlined the concept of stress inoculation and what he called stress inoculation training. A few candidates considered more than one strategy, although the question required one only. In this case all answers were marked and the best answer credited. Some candidates focused on different ways in which stress could be reduced or treated and often wrote about medical and psychological techniques. This strategy could not be credited because prevention (before it happens) is different from treatment (when it has happened).

### Question 8

- (a) This question required candidates to suggest how they would use a fear arousal campaign to promote cycle helmet wearing. Any method could be used as long as fear arousal was a major component. Some candidates suggested conducting a laboratory experiment where one group of participants could be given frightening information about cycling head injuries and another condition of the independent variable for participants who would not be given no such information. The DV could be the number of participants who subsequently wear cycle helmets. Methodological knowledge is needed for a top band mark.
- (b) This question reflected the wording of the syllabus on the health promotion topic, giving candidates a relatively free choice in the one study they chose to describe. Some candidates described the Tapper et al. study which looked at how to encourage children to eat more healthily. Other candidates look at community studies (e.g. Farquhar) which have sought to improve health, specifically heart disease. Marks were allocated according to the detail, accuracy and depth of understanding in the answer.

## **PSYCHOLOGY AND ENVIRONMENT**

Too few candidates to write a meaningful report.

## **PSYCHOLOGY AND ABNORMALITY**

### **Question 13**

- (a) A token economy is a behaviourist strategy used to modify behaviour, to make it more desirable, through the use of positive reinforcers, i.e. tokens which can be exchanged for desired items. Most candidates scored maximum marks usually through competent explanations showing good understanding of the term and through the use of an appropriate supporting example. A few candidates scored no marks because they either did not know the term or because they confused it with something else.
- (b) Most candidates scored some credit, but only a few scored full marks. Most candidates could make the point that if a token is given, the behaviour exhibited might be reduced, but only a few understood how this related to the basics of operant conditioning. Further, only a few candidates could relate it successfully to an impulse control disorder (such as kleptomania or pyromania). A few candidates misunderstood or misread the question and suggested that a token is a reward and a behaviour is more likely to be *repeated* (rather than reduced).

### **Question 14**

- (a) There were two extremes of marks shown in response to this question, with far more candidates scoring higher marks (and maximum marks) than candidates scoring low marks. Answers at the top end of the mark range covered a range of different aspects from the syllabus and usefully brought in the case studies of phobias from the AS part of the syllabus, such as little Hans and little Albert. The case study of Lucy was also commonly included in answers. In relation to treatments for phobias, most candidates wrote about systematic desensitisation and many wrote about blood phobia and applied tension to treat it.
- (b) The evaluation for this question part was often very good and many candidates scored more marks, probably because they were focusing specifically on phobias (unlike the peripheral patient-practitioner relationships as was the case for **Question 6**). Answers at the top end of the mark range included at least three issues and most also considered the named issue of children. Those candidates including the case studies of little Albert, little Hans and Lucy in (a) had an advantage here, although any relevant studies of children could also be evaluated to score equivalent marks. Another common issue included was the debate about competing explanations such as behaviourist learned phobias versus the Freudian psychodynamic explanation.

### **Question 15**

- (a) Candidates were given a free choice of method to use to investigate sex differences in depression. A small number of candidates suggested looking at actual statistics of those so diagnosed, others suggested interviewing practitioners (or sending them a questionnaire) to gather statistics on any differences. A few candidates suggested asking the patients themselves, but this type of answer was often confused when the data collected would be a yes/no for males/females. As always methodological knowledge appropriate to the suggested method was the key to scoring top marks.
- (b) This question asked for a description of evidence claiming there are sex differences in depression. A few candidates made a number of general points, often referring to differences in hormones, but these answers were more anecdotal than psychological. Other candidates organised their answers much more cohesively and considered a wide range of different factors including cultural differences, perceptions by those defining the illness and, of course, differences in hormones.

### Question 16

- (a) Most candidates knew about the work of Frith on ‘cognitive deficits’ and often explained it in good detail. A few candidates knew the basics of his theory and they scored some marks. A small number of candidates did not know the cognitive explanation and so guessed, often suggesting it was ‘cognitive processing that had gone wrong’ for which they could be credited because that is the simple argument Frith proposed, but these were very weak responses.
- (b) This question required an investigation using a laboratory experiment to investigate cognitive deficits as proposed by Frith. Answers at the top end of the mark range included IV, DV, design and controls, and provided a procedure that was easy to follow. Such answers often had participants completing a task where the DV was the number of errors on a task. There were candidates who did not suggest an experiment, so did not answer the question set and the response was not creditable.

### **PSYCHOLOGY AND ORGANISATIONS**

Too few candidates to write a meaningful report.