

# 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, aims, procedure, sampling, results and conclusions. In **Section A**, the candidates' knowledge of aims (**Question 13(a)**), procedure (**Questions 6, 10(a), 14(b)** and **15**), sampling (**Question 4**), results (**Question 12(b)**), ethics (**Question 2**) and evaluation (**Question 13(b)**) was clearly 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 methodology in psychology so that they can see how the study illustrates these principles, for example to be able to explain their answers to **Questions 3** and **5** and by practising responding clearly to the key words in the question, such as 'conclusion' in **Question 10(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.

One final general comment is about handwriting. Examiners' ability to read every candidate's answers is limited by legibility. A script which cannot be read cannot be awarded marks so candidates need to ensure that their answers are legible. One particular problem this year was very small writing. Whilst it is possible to enlarge scripts as they are marked electronically, it is very difficult to read candidates' writing if the letters are smaller than the typeface on the exam paper. Such writing is likely to disadvantage candidates.

## Comments on Specific Questions

### **Section A**

#### **Question 1**

- (a) This question was asking for the meaning of the term, rather than the effects of cognitive load. Many candidates gave simplistic answers in terms of 'overloading the brain' or just defined memory capacity whilst others could only express their answer in terms of lying. However, there were some excellent answers explaining the idea of simultaneous processing.
- (b) The answers here were much better than in part (a), with many candidates opting for the easier route of arguing 'yes' and supplying 'more pauses' and 'less blinking'. Surprising few candidates gave percentages here, yet these are crucial in the study to understanding that many of the results were not supportive of the idea of high cognitive load due to lying producing an identifiable set of behaviours.

#### **Question 2**

There were many good answers especially in relation to ensuring the false story was not psychologically damaging but some candidates gave overly generic answers, such as 'they did not harm them and they debriefed them'. To earn the second mark in each case, the ethical issue needed to be linked directly to Loftus and Pickrell's study. It is sometimes useful to ask students 'Is it obvious from your answer which study you are writing about?' If the answer is 'no', then the response is probably generic. Many candidates also gave inaccurate answers suggesting that the participants gave consent or knew exactly what they were going to have to do (neither of which was the case in this study).

### Question 3

- (a) Few candidates recognised the significance of this part of the study (or knew about it), so were unable to gain many marks. Nevertheless, some were able to do so, making appropriate reference to testing the effects of nurture by attempting to reverse the impairments caused by the experimental treatment.
- (b) The responses to this part of the question suggested that perhaps candidates were unaware of this aspect of the study, as many simply gave partial answers, referring to findings that could equally have been descriptions of the results of the main part of the investigation, e.g. saying that the passive kittens did regain visually-guided paw placement.

### Question 4

- (a) This question was well answered. Most candidates were able to give appropriate examples but there were still some errors, some relating to possible occupations relating specifically to females, such as housewife or waitress.
- (b) Part (b) tended to be answered in very generic terms with most candidates simply saying that it made the study more generalisable or 'representative' – with no reference to obedience at all. One common mistake was to suggest that this was to investigate the effect of job type on obedience.

### Question 5

- (a) Some candidates were aware that the guards had been told that the focus of the study was on the prisoners (i.e. that they had been misled) although many others made a range of errors, for example suggesting that the guards believed they were in a real prison.
- (b) Many candidates were able to correctly identify the importance of keeping the real aim from the guards in order that they did not behave in the way the experimenter expected. However, many such answers were overly generic, simply saying 'So that there would be no demand characteristics', without explaining what this meant in terms of the prison simulation study.

### Question 6

- (a) This question was well answered. The majority of candidates recognised that the question was about the model, so correctly identified the helping behaviour as a key part of their answer. Beyond this, however, fewer were able to describe the behaviour – waiting (to allow real passengers the time to help if they were going to), before acting. Some candidates described the conditions of the IV (e.g. the time differences between conditions) without referring to the behaviours, so could not earn credit. A significant minority of candidates misunderstood the 'model' and described the behaviour of the 'victim' (cane or drunk) so did not earn marks.
- (b) Many candidates earned a mark for simply observing that the model helping caused others to help. However, some then contradicted this by saying the model helping caused others not to help, often this was followed by reference to diffusion of responsibility. Others attempted to make reference to effects on comments made by other passengers. Nevertheless, some candidates who had misunderstood in part (a), and described the behaviour of the victim, recovered here and were able to correctly answer part (b).

### Question 7

- (a) Many candidates were unable to answer this question, suggesting a lack of understanding of this study. In some answers the idea of 'grouping' when present was not always clear and explicit. When candidates did give further detail was typically about over/under estimators (less often with reference to the fact that the groups were in fact random) rather than the alternative of Klee or Kandinsky preferences.
- (b) This question part was answered somewhat better than part (a), with candidates able to identify the importance of the demonstration that the boys chose the maximum difference option.

### Question 8

This question was answered well by many candidates. A minority did not describe the process of identification of aggression level and allocation to groups but described aspects of the general procedure of the study (so did not earn marks). In some cases, candidates described most of the correct procedure well but then said that they groups were made up of children with *different* levels of aggression, i.e. that all the very aggressive children were in one group. Candidates generally seemed unaware that the participants were allocated to three groups (not in 'pairs' to two groups).

### Question 9

- (a) Most candidates were able to gain one mark by identifying Hans's phobia of horses but fewer were able to follow this up with any appropriate detail. A small but significant number misreported that he had a phobia of widdlers.
- (b) Fewer candidates scored marks here, with many suggesting that the benefit was being studied by Freud or resolving his complex. Nevertheless, reference to being able to stay with his mother was seen in some responses although explaining why Hans wanted this was less common. An alternative way to gain marks here would have been to explain that it provided, though the use of defence mechanisms, a socially acceptable release for the expression of his unconscious motives. Again this was seen occasionally, but from even fewer candidates.

### Question 10

- (a) Most candidates were able to gain marks here, with attractive (/unattractive) and black (/white) being most popular.
- (b) Candidates often provided findings here rather than offering conclusions, i.e. they were unable to demonstrate an understanding of what the findings 'meant'. This is an important skill, and candidates would benefit from additional practice of this skill.

### Question 11

- (a) Many candidates seemed to be unfamiliar with the basic findings of this study in relation to incongruent situations (when the motive and outcome did not match), i.e. in this case that the negative valence was the most important element. As a consequence there was a wide range of inappropriate answers with few candidates gaining full marks.
- (b) Although a minority of candidates were aware that this aspect of the procedure made little difference to the children's responses, the majority of candidates appeared to be guessing.

### Question 12

- (a) This question was not well answered, with many candidates unable to describe this aspect of the study.
- (b) Although many candidates were able to answer this correctly and gain the full two marks, there were some common errors. Some candidates explained the difference between the two groups in terms of the information they had been given but did not state the difference in the results. Others muddled the Epi Ign and Epi Inf reactions and still others began their answer correctly, but then made reference to aggression and anger and not euphoria.

### Question 13

- (a) The majority of candidates were able to earn some credit here, typically saying to see if 'we can tell the difference between the sane and the insane'. Many candidates were able to add appropriate detail, such as whether the context of a mental hospital affected this decision-making. Many candidates had misunderstood the purpose of the study, reporting that it was to test whether putting a sane person in an insane place would send them insane. Other candidates demonstrated good recall of the study, providing lots of accurate detail of the procedure, but not the aim of the study, so could not earn marks.

- (b) Responses to part (b) were similar in that candidates often earned some credit, for example by simply saying that it had high ecological validity as it was 'a real place' without mentioning anything specific. Better exam technique is likely to improve such candidates' responses, as they needed to contextualise their answers to gain full marks.

#### Question 14

- (a) This question was generally well answered although candidates often made reference to loss of memory in part (a), but misidentified the initial event, mistakenly describing the letter. A significant number also described the spontaneous appearance of Eve Black rather than the forgotten trip.
- (b) Part (b) was not so well answered as part (a) but nevertheless, candidates showed some knowledge of the study and many were able to report that hypnosis was used to access the different personalities.

#### Question 15

The answers here were generally very good, with many candidates gaining full marks. However some candidates were disadvantaged by a lack of understanding of an independent variable.

#### Section B

#### Question 16

There was a tendency for candidates to give a general evaluation of the study rather than those aspects specifically related to reliability. Some comments earned marks where they were relevant, but in general such candidates scored low marks even though they may have written a great deal. However, if candidates understood the term reliability (and didn't confuse this with validity) they were able to score much higher marks. The strongest responses could discuss in detail two strengths and two weaknesses of the study in terms of reliability.

Although all the studies were used by some candidates, the Demattè et al. study produced the best responses, with references to the 'same' distance from screen, amount of scent, timings, etc. Issues with sample size and characteristics were popular routes for highlighting a lack of reliability.

#### Question 17

In answering this question, candidates' choices were spread over all three studies, with the Dement and Kleitman study the most popular. There was also a wide spread in marks. One common error was to provide lots of detail about what the quantitative data of the chosen study was, but the candidate did not then make links to the strengths or weaknesses of quantitative data. When the question was answered, strengths and weaknesses also needed to be differentiated between the studies, for example – a strength almost always mentioned was 'ability to draw comparisons' across all studies, and inability of qualitative data to understand the 'why' was also stated across all studies. Where candidates did make links to the studies these were often very simple, but often sufficient to raise their mark into a higher band.

# 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 provide answers that equate to mark allocation, so an answer worth two marks should be short and an answer worth ten marks should be correspondingly longer. **Section B** questions are not short-answer.
- For a **Section A** two-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 given to each question part 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

Some candidates confuse basic 'command terms' such as identify, outline, explain and describe. To identify is simply to name something (worth one mark). To identify would provide a term but give no information on what that term means. To outline is to give a brief description (worth one or two marks). To describe or to explain is to give more detail than an outline so the reader has some understanding of what is being described. Explanation and description questions are worth two marks so more detail is required.

## Comments on Specific Questions

### Question 1

- (a)** Many candidates focused on different ethical issues rather than the practical problem caused by the ethical issue of the sensitivity of the video clips, for which they could not be credited. The practical problem that could have been the focus of answers was that because of the sensitive nature of the video clips it was essential that as few people as possible outside the police viewed the tapes which meant testing reliability (of the coders) was a problem.
- (b)** If candidates could not outline the practical problem in part **(a)** then they could not describe how that practical problem was overcome in part **(b)**. The best answer to address the practical problem created by the sensitivity of the clips would be 'the second coder did not see all of the tapes; they only coded a sample of those seen by the first coder (and because the inter-coder reliability was high, the second coder didn't see any more tapes)'.

## Question 2

Most candidates answered this question correctly and many scored full marks. There were a large number of features that could have been included such as scheduling (e.g. the number of interviews and where they were done), the interviewers themselves (e.g. two of them, both female, or aspects of their manner) and various ratings (such as clarity and confidence) and others too. Answers scoring limited credit did no more than identify a feature, whereas answers scoring full marks gave some elaboration or provided a contrast. For example, writing 'rating of clarity' versus 'rating of clarity on a 1–10 scale'. Similarly writing 'interviews were done by telephone' versus 'most interviews were done by telephone but some conducted at the University'.

## Question 3

- (a) A common incorrect response was 'in a laboratory'. Other candidates correctly stated 'in a quiet room', although this scored limited credit without expansion. Elaboration was required to score full marks, and in this instance this was often shown by adding 'in Cambridge or Exeter'.
- (b) This question required an outline of two tasks in addition to the eyes test that participants were required to do. There were three possible answers: (i) the AS/HFA group were asked to judge the gender of the person in each photograph; (ii) The AS/HFA group completed the short WAIS-R intelligence test and (iii) Groups 1, 3 and 4 also completed the Autism Spectrum Quotient Questionnaire (AQ). Many candidates correctly outlined two of these tests and scored full marks. Some candidates could outline only one, and a few candidates could not outline any of these additional tests.

## Question 4

- (a) A small number of candidates misinterpreted this question and described the Milgram study itself, rather than answering the specific question. A few candidates ignored the 'sample' part of the question and stated that the shock ranged from 15 to 450 volts. There were four aspects to the sample shock and any two of the four gained credit: a (single) 45 V shock; applied to the wrist (of the participant/teacher); using the third switch (of the shock generator); from a 45 volt battery (wired into the generator). Alternatively both marks could be gained by elaborating on one of the four possible answers.
- (b) Those candidates stating '15 to 450 volts' in part (a) continued the incorrect answer by writing things like 'to test the level of obedience' showing further misinterpretation of the question. The reason why the participant (the teacher) was given a sample shock was so they believed the situation/apparatus/actual shock was *real*. Also acceptable was that the sample makes it likely to be a *valid test of destructive obedience*. Many candidates wrote about the former belief, but failed to provide elaboration, with no more than 'so it was real' being a typical answer (for some credit).

## Question 5

- (a) Many candidates were able to provide one reason for the prisoner uniform (although a few incorrectly wrote about the guard uniform) but often a second reason was not given, despite the question asking for **two** reasons. The most common reason given to explain why the prisoners were given uniforms was that it created deindividuation. Some candidates provided one reason twice, such as 'reduced their identity' along with 'made them anonymous' when both these are deindividuation. A few candidates referred to the uniform emasculating the prisoners and a few candidates stated that 'it made the situation seem more real' which was also worth credit.
- (b) Many candidates answered this question incorrectly because they failed to answer the question set. The question stated 'For one of these reasons...' meaning the reason provided by the candidate in question part (a). A good example would be to write about emasculation in part (a) and then in part (b) write that 'the smock with no underwear forced the men to sit like a woman'. Another option would be to write 'made anonymous' in part (a) and then 'smock with number – so prisoners didn't use their own name' in part (b). As can be seen from these examples, they also address the second part of (b) which is to *explain how it served that function*.

### Question 6

- (a) Some candidates did not answer the question and instead wrote about the original Bandura et al. study. The question invited candidates to go beyond that study and to suggest a control that could be used in a newly created/hypothetical study. Any sensible control received credit and for maximum marks there needed to be some explanation of the reason for the control, however brief. Most candidates focused on the appearance of the models: for example, the aggressive model mustn't look nastier (or nicer) than the feared model; others looked at familiarity: the feared model should not be someone well-known unless the other model is too.
- (b) Candidates should always aim to provide some elaboration in answers that begin with 'Describe'. Many candidates simply identified a guideline, e.g. 'right to withdraw', without elaboration, for limited credit. Brief elaboration of how 'right to withdraw' applies to the proposed study would have gained these candidates additional credit. For example, '*right to withdraw* such as letting the children know they can leave at any time, if they become frightened'.

### Question 7

- (a) The question required candidates to describe the results for good/bad motive only. The result of any other variable scored no marks. There were two different types of answers candidates provided, both scoring maximum marks. First was that three-year-olds, just like seven-year-olds make use of motive information to make moral judgements. The second type involved the comparison of mean scores; such as good motive/outcome were 6.55 for three-year-olds compared to 6.20 for seven-year-olds. Answers did not necessarily need data for two marks, but it was one way to add detail.
- (b) This question part asked about the results for the implicit/explicit motive. Although there were a few perfect answers, many candidates were confused between how this variable combines with good and bad motive and good and bad outcome. A perfect answer would be: children are more likely to judge by motive when it is explicit, though only for bad motive with good outcome and good motive with bad outcome, i.e. when there is a conflict between motive and outcome (non-congruence).

### Question 8

- (a) A small number of candidates incorrectly wrote about the questionnaire used to *create* anger, and a few candidates incorrectly wrote about the questionnaire that used a rating scale (which gathered quantitative rather than qualitative data). Other candidates provided partial answers when writing about, for example, 'the open ended questionnaire' and nothing more. Elaboration was needed, e.g. 'an open ended questionnaire asking participants to describe physical/emotional sensations experienced during the experiment'.
- (b) Nearly all candidates scored some credit when giving one strength of qualitative data. Typical answers were 'the participant's answer can have a lot of detail'. What was often missing was the elaboration, as required by the second part of the question, which needed candidates to contextualise the strength *in this study*. For example, if a candidate wrote 'it allows the participant to give a true or valid answer about their emotional state' then it is clear from this answer that the answer is related to this study.

### Question 9

- (a) A significant number of candidates would have benefitted from a more careful reading of this question. For this question part many candidates failed to answer the former part of the question (asking for a description of what Dement and Kleitman were expecting to find) instead answering only the latter part of the question (stating what was found). The simplest full answer would be to write 'they were expecting to find a positive correlation and they found one.'
- (b) Some candidates had explained the results in part (a) and so wrote the same answer again. Reading both parts of the question before responding would help avoid this common error. The simplest full explanation for the results was that the longer a person spends in REM sleep, the longer the dream will be so there is more detail to recall about it.

### Question 10

- (a) The question focuses on the control task that was compared to a routes task rather than any other task the candidates had to perform. Some candidates gained limited credit (usually for outlining the control task) and some candidates correctly outlined both tasks for full marks. The strong answer would be: the control task involved speech output, from repeating (two four-digit) numbers, whilst the routes task involved a topographical sequencing task, with participants describing a route between two points (in London).
- (b) Candidates had to identify any two areas of the brain that were associated with the routes task activation, and any two from the following list were correct: extrastriate cortex, medial parietal lobe, posterior cingulate cortex, parahippocampal gyrus and right hippocampus. A number of candidates stated 'the hippocampus' which was too general.

### Question 11

- (a) Many candidates were not able to describe counterbalancing and either did not respond, or appeared to guess an answer, which nearly always scored no marks. Some candidates explained the purpose of counterbalancing, which is to reduce order effects, but did not write what counterbalancing actually is. Those candidates scoring full marks explained that counterbalancing is presenting each possible order of the conditions of the IV to different participants, e.g. if there are two conditions, the order for participants is alternately A then B and then B then A.
- (b) Like **Question 8(b)**, this was an 'in this study' question. Marks were available for explaining why counterbalancing was necessary (which is to avoid order effects (in a repeated measures design)) and for elaboration in the context of this study (so in this experiment they didn't have the effect of one smell influencing their judgment of attractiveness on the next trial), and answers scoring full marks covered both elements.

### Question 12

- (a) Reading both parts of the question before responding would help avoid a common error, that answers to part (a) often included behaviours that belonged in part (b). A few candidates incorrectly wrote about gaining entry to the mental institution but most candidates scored partial or full credit for writing that pseudo-patients were instructed to demonstrate to staff that they were sane.
- (b) Any two behaviours received credit here, but these must have been behaviours that showed the pseudo-patients were sane and would therefore contribute to their discharge. For example they (pretended to) take the drugs they were given; joined in with hospital activities; and were friendly/co-operative/exhibited no abnormal behaviours. Answers that were not acceptable were those that did not contribute to discharge, including queuing for food (said to be 'oral acquisitive syndrome') and note-taking ('patient engages in obsessive writing behaviour').

### Question 13

- (a) Most candidates stated that Eve White had a good relationship with her parents and Eve Black a poor relationship. This is partially correct, because Eve Black did have a poor relationship due to a 'strong feeling of rejection by her parents, especially after the birth of her twin sisters'. Eve White however, admitted 'difficulty in her relationship with her mother' and being bewildered by punishments for misdemeanours she could not recall, which could not be accurately described as a good relationship.
- (b) There were a number of different ways in which Thigpen and Cleckley obtained information about the relationship between the Eves and her parents. They used hypnosis but the main source was 'over 100 hours' of interviews with both Eves. Thigpen and Cleckley also interviewed Eve's parents.



#### Question 14

- (a) A number of candidates provided no answer at all in response to this question, despite the fact that empathising is a major feature of this core study. Billington herself defines empathising: “The affective component of empathising involves an emotional response that arises as a result of the comprehension of another individual’s emotional state”. Whilst candidates never need to learn specific definitions they do need to be able to describe what is meant by a term in their own words. In this instance, writing about ‘an emotional response’ would gain some credit and writing about ‘understanding someone else’s emotional state’ would gain full credit. *Identifying* an emotion in a person is not enough; it is the *understanding* of that emotion that typifies empathising.
- (b) Candidates needed to identify two different features of extreme empathisers, so by ‘female’ and ‘humanities students’ scored full marks. Alternatively candidates could expand on one feature such as identifying that they were ‘humanities students’ for one mark and then going on to give more information, i.e. elaboration such as ‘2% were from humanities, 0.5% from physical sciences’.

#### Question 15

The words *focus of attention* appeared both in the stem and in the question to emphasise the importance of it to the question. Many candidates did not acknowledge this and instead gave two differences between the BDD patients and the controls that were not related to the focus of attention, and would have benefitted from more careful reading of the question. A number of candidates gave just one difference, but two were required. Limited credit was given for identification and elaboration was required for full credit. A correct answer would be: BDD patients were more likely to focus on an internal impression or feeling than their external reflection in the mirror, common in control participants. Alternatively for the second mark candidates could have given some data, such as ‘BDD patients  $-0.49$ , controls  $-2.2$ ’. The second difference relating to focus of attention was that those with BDD were more likely to focus on specific parts rather than their whole appearance.

#### Question 16

This question, which allowed any weakness to be included, resulted in a number of answers scoring the maximum mark. Other candidates who didn’t have quite as much detail or quality in their answer still scored high marks. In the middle range a number of candidates had pre-prepared weaknesses and these were applied to a named study whether they were relevant or not. Sometimes answers showed a lack of understanding of the study itself. Weaker responses showed three types of limitations: (i) writing about all three named studies rather than just the named study. In such cases all answers were marked and the best one credited, but it often meant that there was much less detail on any one study. (ii) Candidates who only considered one weakness and often wrote anecdotally, and (iii) candidates who included very little psychology and scored few or no marks. The study by Freud was chosen by most candidates and weaknesses included little Hans being a child and associated ethical issues, the use of leading questions; the bias on the part of the father and of Freud; and the lack of scientific objectivity and lack of quantitative data.

#### Question 17

Answers for this essay question were not as strong as those for **Question 16**. Discussion must include both strengths and weaknesses. Some candidates struggled to provide both strengths and weaknesses, any answer that provides only strengths or weaknesses will gain limited credit. Although there can be some imbalance in the length of a strength or a weakness, there should still be at least two of each, providing one strength and three weaknesses is imbalanced and will gain limited credit. In addition to these technique errors, there were also errors when relating the strength or weakness to the named study and candidates are encouraged to think carefully about what they write. Pre-prepared answers are of little value if they cannot be related to one of the named studies.

# 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 and conclusions. In **Section A**, the candidates' knowledge of background (**Questions 7(a) and (b)**), apparatus (**Question 2(a)**), procedure (**Questions 3(a) and 12(a)**), results (**Questions 3(b), 11(a) and (b)**), conclusions (**Question 12(a)**) evaluation (**Questions 5(a) and (b)**) was clearly 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 **Questions 1 and 14**, and of the procedure in some cases (e.g. **Questions 4 and 10**). 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 explain their answers to **Questions 1(b) and 15**.

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.

One final general comment is about handwriting. Examiners' ability to read every candidate's answers is limited by legibility. A script which cannot be read cannot be awarded marks so candidates need to ensure that their answers are legible. One particular problem this year was very small writing. Whilst it is possible to enlarge scripts as they are marked electronically, it is very difficult to read candidates' writing if the letters are smaller than the typeface on the exam paper. Such writing is likely to disadvantage candidates.

## Comments on Specific Questions

### **Section A**

#### **Question 1**

- (a) Candidates needed to consider an innocent suspect being questioned about a crime (that they therefore had *not* committed), also in a 'high stakes' situation. They risk wrongly losing their freedom if they cannot convince a court that they are not guilty. This issue is raised in the study itself, and the candidates were required to suggest why investigating such individuals would be useful – simply because the two might be indistinguishable if the effect of a high stakes situation is the same. Answers largely focused on controlling lying, which was not relevant to the question.
- (b) Some candidates were able to score marks here, suggesting factors to control such as the nature of the crime.

#### **Question 2**

- (a) This question part was very well answered, with many candidates describing the black and striped walls and some adding detail such as that the stripes were vertical.
- (b) Candidates were less confident here, with confused answers about the pattern preventing the kittens from seeing each other or to create an *illusion* of depth. There were some good answers relating to controls, for example suggesting that the ensured that the only difference in visual experience of the active and passive kittens was due to their ability or inability to move independently.

#### **Question 3**

- (a) This question was well answered, with candidates typically being able to describe either the observer asking passengers questions or recording spontaneous comments, or both.
- (b) This question was also well answered, with many candidates being able to give examples of comments from the passengers.

#### Question 4

- (a) Many candidates scored limited credit, typically for 'doll', but there were many confusions with toys such as the Bobo doll or Tinker toys that could not earn credit. It is important that candidates understand the different stages of this study.
- (b) Although a small number of candidates demonstrated a clear understanding of the function of the aggression-instigation phase of the study, the majority did not. Candidates typically responded, incorrectly, with the idea that the children were being tested.

#### Question 5

- (a) Many candidates answered this question well, suggesting an appropriate advantage, such as children being less likely to be aware of the aims than adults or to observe developmental change. However, many candidates did not respond to the 'in this study' instruction in the question so did not make their answer relevant to Freud's study of little Hans. It is important that candidates are aware that that this instruction in a question means that they must relate their response to the study rather than just giving a generic answer. Those candidates who did answer the question in full often provided succinct links, such as 'Children are less aware of the aims... like little Hans did not know his fantasies were being studied' or 'Using children can help in studying development... so it was easier for Freud to look at the effects of the Oedipus complex on little Hans than on his adult patients.'
- (b) Again many candidates answered this question well, often in relation to problems with children being easily misled. In common with part (a), some candidates only gave generic answers when they could easily have linked their answer to the study. For example 'Children are easily misled... like little Hans when he was asked leading questions'. Another common way to earn marks here was to refer to the ethical problems with studying children, such as the difficulty of obtaining valid consent.

#### Question 6

- (a) This question was not well answered. Many responses appeared to be guesses, such as saying babies preferred female faces.
- (b) This question was also not well answered. Many responses here too appeared to be guesses, such as saying babies preferred black or preferred white faces (even though no comparison had been made in the study).

#### Question 7

- (a) There were many good answers from candidates here, explaining 'outcome' in terms of the alternative endings to the story used in the study. Weaker responses often earned a mark for 'common sense' explanations of the meaning of 'outcome' without relating this to the study or to children's moral understanding.
- (b) Although there were some good responses here, this question part was somewhat less well answered than part (a), with responses typically suggesting a limited or muddled reason.

### Question 8

- (a) Most candidates were unable to answer this question, with many ignoring the part of the stem referring to self-report scales. Such answers focused on the physiological changes. Other candidates described the behavioural and emotional changes in the participants, such as their engagement with the stooge.
- (b) Some candidates were able to gain marks here by making points about strengths and/or weaknesses of self reports in general that happened to be relevant to the scales used.

### Question 9

- (a) Many candidates appeared to be guessing answers here, suggesting that participants slept for an average of 8 hours per night. Some students who did demonstrate specific knowledge of the study mistakenly quoted findings relating to time spent in REM rather than time spent asleep. These are two different pieces of information from the study.
- (b) Responses were slightly better in part (b) of this question, with candidates offering sensible ideas about normally sleeping for longer, for example as they would be less comfortable sleeping in a different bed. Few made reference to the fact that they were frequently woken, suggesting that some candidates did not know basic details about the study.

### Question 10

- (a) Many candidates were able to suggest that the tasks related to describing famous landmarks, but some thought that these were all in London (they were specifically *not* in London).
- (b) In response to part (b), many candidates were able to identify the routes task as an appropriate comparison, although fewer were able to explain that both were topographical tasks (but that landmarks information was non-sequencing whereas routes information was a sequencing task).

### Question 11

- (a) Most candidates were able to give a brief answer, observing that participants found faces more attractive when paired with a pleasant smell than with an unpleasant one. However, answers often lacked any more information than this, rarely providing examples of pairings and even more rarely offering accurate data.
- (b) Fewer candidates were confident in their answers here, with some confusion over the nature of the control condition and of the results. Better answers provided both information about the minimal (non-significant) difference and that this was between the pleasant smells and clean air.

### Question 12

- (a) This question part was well answered, although responses were varied. In addition to ideas on the mark scheme, candidates suggested that arriving at a hospital and checking yourself in is unusual, or simply that the pseudo-patients reported their symptoms convincingly.
- (b) This question part was not so well answered as part (a). Some candidates offered acceptable explanations arguing that the reason given was 'wrong' in the moral sense.

### Question 13

- (a) This question was not answered well, with few describing the context of the event.
- (b) Even candidates who had not earned marks in part (a) were sometimes able to earn marks here as they recognised the importance of Eve Black's tendency to lie, or her general attitude to Eve White's child, i.e. one of annoyance. This enabled candidates to suggest a plausible response for which they could earn some credit.

#### Question 14

- (a) There were a range of appropriate answers to this question, including the technical and spatial domains. However, candidates were often unable to answer this question. It is important that candidates have a basic knowledge of the underlying principles being explored in each study.
- (b) Very few candidates were able to answer this question. Since the concept of 'systemising' is central to this study, the meaning of a 'system' is important in the understanding of the concept of systemising itself.

#### Question 15

- (a) This question part was typically well answered, with candidates typically identifying the presence of an IV and DV and/or controls with better candidates able to define IV and DV.
- (b) Although candidates could often identify a generic advantage of experiments, such as controls or manipulation of the IV, fewer were able to relate this to Veale and Riley's experiment as required by the question. This suggests that, although they had learned the basic notions of what an experiment is, this understanding was not sufficiently well developed to be able to apply it, even to a known study.

#### Section B

#### Question 16

In answering this question, candidates' choices were spread over all three studies, with all producing some good answers. Candidates were able to focus their answers on strengths although often this was superficial. Some were able to effectively illustrate the strengths they had identified with examples from their chosen study, for example giving examples of control groups from the study and, in the best cases, saying what these groups controlled for. In Baron-Cohen et al.'s study, for example, candidates identified having IQ matched controls to ensure that lower scores on the eyes test produced by the HFA/AS group were not simply due to lower ability. Some candidates spent a lot of time describing the study, which was not creditworthy; it is important that candidates recognise that in these essays description of the study cannot earn credit as the question is asking for evaluation.

#### Question 17

The most popular choice here was the Milgram study. The quality of responses was varied although many were able to identify basic elements of situational factors such as the presence of the authority figure and the location in a university. However, the answers typically earned only limited credit, as they gave few or only brief examples of situational factors. For example, mentions of the prods were remarkably rare as was reference to the experimenter's clothing. Very few candidates attempted to argue against the situational explanation, which they could have done by referring to the individual differences between participants or the percentage *not* reaching 450 V. Some candidates did make reference to payment but incorrectly identified it as a situational factor when in fact Milgram suggested that this was not an important influence as when a student sample were given no payment, they still obeyed.

# 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 identify and/or define the various types of observations in part **(a)**. Candidates should suggest a simple alternative to the original study in part **(b)** and give clear details of the procedure followed. Extended evaluative points that make direct reference to the alternative idea are necessary in part **(c)** to achieve full marks.

#### **Question 2**

It is also important for this part of the exam that candidates are aware of all of the issues listed in the syllabus. A minority of candidates did not know what control and/or reductionism meant and therefore did very poorly in the parts of the question referring to these issues. In addition, 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 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 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 wrote 'introductions' before beginning to answer the questions, which is not a good use of the time available.

A significant minority of candidates answered both questions in the **Section B** essay. When a candidate did this they were awarded the mark for the best of the two questions (**Question 3** or **Question 4**). These candidates usually achieved very poorly. It appeared that this rubric error was more common compared to previous sessions.

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 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 participant/non-participant, natural and controlled observation and overt/covert. Many just listed terminology which achieved limited marks. Some just described what happened in the Zimbardo study, although they were given limited credit as this was a description of one type of observation.
- (b) The vast majority of candidates were able to provide an alternative way to investigate social roles and some included sufficient detail to be fully replicable and were awarded marks in the top band. Many examples of teacher/candidate interaction were provided as well as observing guard/prison interaction in real prisons. There were some more innovative designs relating to the workplace, restaurant settings, and public services e.g. military and hospitals. There was a tendency to not include enough detail on how the behaviours were being recorded i.e. just 'observation'. If they did not have the knowledge of how different observations take place in different settings, then they would not be able to fully describe an observational study in enough detail to gain higher marks. Likewise some candidates responded to this question by just describing or sometimes evaluating Zimbardo's study, although this was limited to a few responses. Some candidates tried to reproduce the study by Piliavin and Asch but this was not done well, this may be because they recognised the social part of social roles but were not able to design an appropriate study. There was good evidence again in the use of anagrams and plans to help candidates recall the key sections to be included in their answers but there were also times when candidates evaluated their study as they described it, which was unnecessary. Some candidates also compared their design to that of Zimbardo et al. as they described each section, which was again not required.

Many did not include the other details required such as where the study would take place, who the participants would be and the duration of the study.

In addition, the candidates need to ensure the 'what' and 'how' for the procedure are very clear.

Some candidates evaluated their idea in this question, for which no additional credit is available. Evaluation is not required in part (b), but is required in part (c).

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

Many discussed issues about the ethics of studying participants who are under 18 (as many studies were done in schools) as well as discussing the ecological validity of the alternative idea. Many also referred to the ethical issues of observing participants.

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

#### Question 2

- (a) The majority of candidates achieved at least one mark for their response to this question. Most were able to give a very brief definition of control and were aware that this was something that was kept standard and/or constant in the study. Many did refer to the control of extraneous variables but the vast majority were unable to give a full definition of what is meant by a 'control'. Some did achieve full marks and had a good understanding of the term.
- (b) Most candidates achieved one or two marks for their responses. Candidates referred to the materials used in the study (e.g. eyes test, IQ and AQ tests) but just named them rather than describing the test fully. Few were able to achieve full marks as they could not say why the control showed standardisation or consistency in the study. A few also did refer to the control group of 'normal' participants but most were unable to explain why this acted as a control.

- (c) The vast majority of candidates did achieve some marks in this section. Most had a basic understanding of the term reductionism and could apply this to the Baron-Cohen et al. study. Popular points included discussing the limited nature of the testing of participants and the limited sample. A few did explain why the study was more holistic due to the different types of participants that were tested as well as the more complex eyes test that was used in this study. Many had clearly been taught the incorrect Baron-Cohen et al. study as they referred to the two choices of words (rather than four) and the participants with Tourette's syndrome who were not in the specified study.
- (d) Most candidates were able to achieve some marks for this question. Many could give both strengths and weaknesses of reductionist research. Popular points included focus on one issue, good reliability, poor ecological validity, missing other factors and a more limited sample. Many of these points were then linked to the Baron-Cohen et al. study. The links were often quite brief and sometimes it was just a mention of the study that was not backing up their previous point. Candidates needed to provide 4 points (two strengths and two weaknesses) all linked to the study to achieve higher marks for this question.

## **Section B**

### **Question 3**

- (a) The vast majority of candidates achieved limited marks for this question by stating it was something to do with everyday life. Many could achieve full marks by extending their answer. A few tried to link their response to the generalisability of the sample which was incorrect.
- (b) Few candidates achieved full marks for this question. Many were able to describe a feature that lacked ecological validity although this was often discussed very briefly and achieved lower marks. Some referred to what might happen in everyday life and this helped them to achieve more marks. Popular responses included focusing on the story told in the Nelson study as well as the lack of realism of the rating scale. Candidates were able to explain how the smells and faces in the Demattè et al. study lacked ecological validity. Responses on the Schachter and Singer study were not well answered, candidates often just stated it was done in a laboratory environment. Some were aware that the injections were unrealistic.
- (c) Most candidates were unable to identify three advantages of ecologically valid research and were not always able to link their response to any evidence. Popular responses included how this made studies more realistic, generalisable and also useful. Many tried to explain how this also made the studies more reliable and representative but this was incorrect.

### **Question 4**

- (a) Many who attempted this question achieved limited marks for stating that ethnocentric bias is where a limited range of participants are used from one ethnic group. Many did give examples but most did not focus on the bias part of the definition. A few candidates were well prepared and described how ethnocentric bias is favouring your own ethnic group over another.
- (b) Candidates achieved well on this part of the question. Most were able to achieve some marks for each study. Many provided their strongest answers describing how the data were collected in the Piliavin et al. study and described very clearly the data collected by the two observers in the study. The Tajfel and Rosenhan studies were not described in as much detail and often achieved fewer marks. Candidates were not able to accurately describe the matrices used in the Tajfel study.
- (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 such as the issues with generalisability and bias in the studies. Some wrote about the lack of generalisability of all three studies but achieved limited marks as this was just one point.



# PSYCHOLOGY

---

Paper 9698/22  
Core Studies 2

## Key Messages

### **Section A**

#### **Question 1**

Candidates need to know what experimental (participant) designs are and how they relate to the study in question. They need to suggest simple alternatives to the original study in **part (b)** covering what, how, who, where and when. Extended evaluative points linked to their own study from **part (b)** are necessary in **part (c)** to gain full marks. There were some examples of unethical studies for **1(b)**.

#### **Question 2**

It is important for candidates to know how each study is linked to the methodology and data presentation, so for this examination, how the nature-nurture debate may be linked to the Billington et al. study. For **part (b)** candidates need to explain how the Billington et al. study could be nature or nurture. For **part (c)** candidates need to evaluate the individual differences approach using the Billington et al. study as an example throughout, rather than just evaluate of the study. Also, to gain high marks, candidates need to write about two strengths and two weaknesses as a minimum. For **part (d)** candidates need to be able to explicitly compare the individual differences approach to any other different approach using studies as examples of their comparison points.

### **Section B**

Candidates must focus their answers in **part (b)** to what feature(s) the question is asking (in this exam how data were collected or how each study was valid) rather than just writing in general about the study. Candidates need to make three separate points in **part (c)** and have evidence from studies for each to gain full marks.

## General Comments

The marks achieved by the candidates sitting this examination covered the range of possible scores but with only a few gaining top-end marks. Some candidates provided a range of excellent answers to many of the questions and could explain psychological terminology well, showing they had prepared themselves well for this paper.

Time management appeared to be good for the majority of candidates. There was some evidence that candidates who over-answered **Question 3(b) or 4(b)** wrote much shorter answers for **3(c) or 4(c)** as a result. Candidates need to ensure they have enough time to answer all questions to the best of their ability.

Candidates need to be aware that they need to answer **one** of the two questions for **Section B**. When a candidate did answer both questions, both were marked and they were awarded the best mark (**Question 3 or Question 4**). These candidates usually achieved poorly.

Candidates need to cover the entire syllabus so that they can respond to both **Questions 1 and 2** as there is no choice with these questions. In addition to this, candidates must include evidence in **part (c)** of their **Section B** essays to achieve the higher marks available. **Question 3** was more popular than **Question 4** and tended to be answered to a higher standard.

## Comments on Specific Questions

### Section A

#### Question 1

- (a) Many candidates could describe what an independent groups experimental design was but there were some who outlined repeated measures. The IV of ill/drunk was cited the most as an example followed by race of victim and then position of model for the Piliavin et al. study. Some candidates simply described the study and scored little or no credit.
- (b) There was a wide variety of ideas given by candidates on how to examine bystander behaviour in a laboratory. Many candidates could appropriately choose a sample and outline an appropriate sampling technique. Candidates usually did well outlining the *what* (the scenario that the participant had to go through) and the *where* (e.g. a university laboratory). Some candidates did not tackle the *how* clearly (the actual recording of the behaviours shown by the participants; simply writing 'behaviours were recorded' was seen as a major omission. There were a minority of highly unethical studies (usually very intrusive imagery or scenarios being used with the participants) presented in this answer that only scored very limited marks – candidates must think like a professional psychologist when designing these studies.
- (c) Many candidates could highlight one or two evaluative points about their own study they had designed in **Question 1(b)**. Common points made were about the sample used, the unethical nature of a study about bystander behaviour and psychological stress, and the validity of laboratory based studies. Many candidates made a series of brief points linked to their own design to gain more marks. Some candidates evaluated aspects generically, but gained limited credit as they had not linked these points specifically to their own study.

#### Question 2

- (a) Many candidates clearly knew what the nature-nurture debate was and gained full marks. However, there were a number of candidates who did not state which explanation was nature and which was nurture in their answers.
- (b) Many candidates could not explain how the Billington et al. study supported the nature or the nurture approach. They simply wrote out the results of the study without using it to state how it supported either side of the debate. Those who did used a genetic argument or a socialisation argument to gain credit.
- (c) Many candidates attempted this question and gained some credit. There was evidence of a minority of candidates evaluating the Billington et al. study in general. These responses could only gain credit if the answer was linked to clear points about the individual differences approach in general. Many candidates could give some strengths and weaknesses of the individual differences approach but then failed to use the Billington et al. study as examples of these strengths and weaknesses so gained partial credit. Some candidates did not follow the instruction '...using the Billington et al. study as an example', and would benefit from a more careful reading of the question set.
- (d) A minority of candidates gave excellent answers comparing the individual approach to another approach using examples of studies throughout – this made the comparisons explicit and high marks were gained. Many candidates were able to describe the individual differences approach and one other approach in some detail but as separate paragraphs with no direct comparisons. These answers could only gain partial credit. Many candidates simply compared two different studies from opposing approaches, without comparing the *approaches*. The question required the candidate to focus on the approaches using studies as examples.

### Question 3

- (a) Many candidates could make reference to at least one aspect of the developmental approach (e.g. behaviour changing) with an example to score full marks.
- (b) There were many very good answers to this question as candidates could pick out the necessary aspects of each study that showed how data were collected. Specific details about each study in terms of the actual data were crucial to gain marks for each study. The Langlois et al. study tended to be the study where candidates obtained the least marks. They tended to focus more on the initial set up of the study rather than the data collection from infants. Many candidates could name at least two of the measures taken in the Nelson study, showing good knowledge of the study. The Freud study was covered well too, with many candidates being able to pick out the different ways in which he collected data about little Hans. There were candidates who gave very long answers here that covered all of the study rather than focusing on what the question is asking – in this case, data collection – and this is not a good use of time.
- (c) Some candidates could only manage brief answers here which could indicate they were not well prepared or that they had run out of time to write a more detailed response. Many candidates could at least outline some advantages like longitudinal aspects of research and usefulness and some then used a study to elaborate on the advantage. However, only a few candidates then went on to relate studies to all advantages. Some candidates made the same point repeatedly, which cannot be credited more than once.

### Question 4

- (a) Some candidates were able to outline the meaning of the term validity, and gave examples. However, many could not show understanding of the term, and wrote about reliability which could not be credited.
- (b) Candidates appeared to know the three studies well but not in the context of validity. The Haney, Banks and Zimbardo study was covered well in terms of the procedure but it was rare to see a candidate link this to validity. The same applied to both the Veale and Riley and Loftus and Pickrell studies. There were candidates who gave very long answers here that covered all of the study rather than focusing on what the question is asking – in this case, validity – and this is not a good use of time. Finally, there were occasions when a candidate argued why the study was *not* valid which is not what the question required.
- (c) Many candidates could only manage brief answers here which could indicate they were not well prepared or that they had run out of time to write a more detailed response. Many candidates could outline one or two brief problems and this tended to be about difficulties making the study valid or ethics. Only a minority of candidates could make three separate points and fewer could relate all to a study in order to gain maximum marks. As with **3(c)**, some candidates made the same point several times using different studies (especially about ethics) but this could still only score limited marks, for one well-made point with evidence.

# PSYCHOLOGY

---

Paper 9698/23  
Core Studies 2

## KEY MESSAGES

### **Section A**

#### **Question 1**

Candidates should suggest a simple alternative to the original study in part **(b)**. In addition, they must ensure they follow the instructions in the question as it asks for a field experiment and some did a laboratory experiment. 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 students 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 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

There was a very small entry for this version of the paper. The marks achieved by candidates sitting this examination did varied across the mark range. Some 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 all were able to finish the paper.

Only one or two candidates answered both questions in the **Section B** essay. When a candidate did this they were awarded the mark for the best 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 4** was the more popular choice of question.

## **Comments on Specific Questions**

### **Section A**

#### **Question 1**

- (a) Most candidates did quite well for this question and were able to describe both qualitative and quantitative data. Many could also give examples of how this data might be collected (e.g. open/closed questions, observations, etc.). No additional credit was available for strengths and weaknesses of the types of data, as this is a 'describe' question.
- (b) Most candidates correctly described a field experiment although a small minority chose to conduct their experiment in a laboratory.

There were a variety of different ideas given as a field experiment alternative for the Schachter and Singer study. Many gave the injections of adrenaline prior to the study but some had the participants take this in an unexpected way (in a drink or at the doctor's surgery) and then observed the participants' behaviour. Some collected the results via self report at the end of the study. Many included stooges to try and influence the behaviour of the participants.

Most 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. A tally chart of observed behaviour would have been helpful or the questions asked of the participants if it was a self report that was used to collect the data.

Some candidates evaluated their idea in this question, for which no additional credit is available. Evaluation is not required in part (b), but is required in part (c).

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

Many discussed ethical issues, problems of administering the epinephrine as well as issues with the data collected. Some also discussed generalisability of the sample.

A few gave well developed points that achieved very high marks as they referred directly to their alternative idea. A few 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. They could give an example of an ethical guideline and many were able to describe that it was a set of rules that psychologists must follow when carrying out research.
- (b) The vast majority of candidates did well on this question and referred to how ethical guidelines were followed (e.g. some participants were allowed to leave the study). A small minority referred to an ethical guideline that was broken in the study.
- (c) Most candidates achieved some marks in this section. They were able to explain why ethical guidelines need to be broken (to add to the ecological validity or to prevent all participants from just withdrawing from the study) and also the problems with breaking ethical guidelines (the harm to participants and how it lowers the status of psychology). Many were aware of the plural nature of this question but were not always able to provide two strengths and two weaknesses to achieve the higher marks. Most were able to put their answers into the context of the Haney, Banks and Zimbardo study.
- (d) There was a mixture of responses to this question. Some candidates did not know what the term determinism meant so did poorly on the question. Many did know what it meant and did quite well and were able to argue that the Haney, Banks and Zimbardo study can be viewed as both deterministic as well as on the free will side of the debate. Those that did address the question were able to put their response into the context of this study.

**Section B**

**Question 3**

- (a) Most candidates were able to give a brief definition of the individual differences approach and achieved at least some credit.
- (b) Candidates often focused for this question on how the data were collected rather than on the behaviour that was investigated.
- (c) Many candidates could identify at least one problem of investigating the individual differences approach. Many focused on the issues of generalisability as well as ethical issues. Some were able to identify two or three ideas and gave some evidence to back up their points.

**Question 4**

- (a) Candidates were able to define what is meant by quantitative data although a few simply said it was numerical data and achieved limited credit for their response.
- (b) The marks for this question covered the range of the mark scheme. Most were able to access some marks by giving a brief description of the study. Many could describe how the quantitative data were collected and gave a number of details that achieved higher marks. Some described how the qualitative data were collected and which could not be credited.
- (c) For this question, candidates need to identify and discuss three advantages with clear reference to a core study for each point. Many were able to describe one or two advantages such as the issues with being able to make comparisons of groups as well as do statistical analysis. A few then described how this could make the research more useful. Many did not link their responses to a study and achieved fewer marks. A few gave clear links to research and achieved high marks.

# PSYCHOLOGY

---

Paper 9698/31  
Specialist Choices

## Key Messages

- Candidates should provide answers that equate to mark allocation, so an answer worth two marks should be short and an answer worth eight 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 four sides of paper in length.
- Candidates should read all parts of a question before beginning to respond 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 just from what has been learned for Paper 3), to 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 four marks, the allocation of marks is 2 + 2, whereas if a description of one study is required for four marks, then the same amount in total should be written as for the 2 + 2 format. Sometimes candidates wrote far too much for two marks, particularly evident in **Questions 1(b), 5(b) and 9(b)**.
- There were some excellent answers for question part **(b)** that were relevant, accurate, detailed and most showed good understanding.

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

The quality of many part **(a)** answers was very good with many candidates using what is included in the syllabus, covering a wide range of sub-topic areas and describing studies well and in detail.

An important message, for question part **(b)**, which is often repeated, is that candidates need to ensure that they know the difference between 'describe' and 'evaluate'. **Section B** question part **(a)** is 'describe' and question part **(b)** is 'evaluate'. 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 (who also gain limited marks);
- those who do not evaluate at all (and score no marks) merely describing more information.

It appeared that some centres had advised candidates to use the same four evaluation issues for every question. This strategy is not recommended, as it meant that candidates wrote about issues that did not apply to the question. For example, a candidate might write 'pain is high in ecological validity' or 'organisations are ethical' both of which have very little meaning. There are many issues that can be applied to every topic area and candidates are advised to think about and choose issues appropriate to the topic area of the question.

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

One question part invites a candidate to describe and 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 of what 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 requires candidates 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 often use the term experiment incorrectly when they are conducting an observation or questionnaire.

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) Most candidates scored full marks for this question part. The mark scheme required two different aspects of bullying to be included in the explanation such as actions, reasons or examples and nearly all candidates managed to achieve this. One explanation is that bullying is a distinctive pattern of deliberately harming and/or humiliating another person.
- (b) Nearly all candidates gave two explanations for bullying and nearly all candidates scored some credit for each answer. What was often missing was the detail in the answer needed for full marks. Often a candidate could give an explanation, such as 'children bully to seek attention', but there was no additional detail to explain what this meant or there was no example given in support. The range of explanations provided by candidates was vast and any appropriate explanation received credit.

##### **Question 2**

- (a) There was a significant minority of candidates who only considered learning styles, not writing about teaching styles at all. These candidates gained limited marks because they only answered half of the question. At the top end of the mark range candidates often included the learning style models of Curry and Grasha and the types of teaching styles proposed by Bennett and Fontana. Measures of styles were less commonly included but ways in which learning could be made more effective was often included and details of it were competently described.
- (b) The named evaluation issue here was 'the reliability and validity of measures' and this issue has appeared in other Education questions because it is a very important issue for this option. Some candidates addressed the issue well and wrote impressive answers. Some candidates confused the terms and some candidates did not attempt the issue at all. To confirm: reliability is the consistency of the measurement. Reliability of a questionnaire/test (e.g. ASI or Kolb's styles) can best be tested using test-retest. Validity is whether a test measures what it claims to measure: learning styles should measure learning styles. A reminder that this named issue is just one of the advised minimum three issues an answer should include.



### Question 3

- (a) Many answers were anecdotal and needed to show evidence of the psychology learned. The purpose of this question type is to test application skills, where candidates apply what they have learned to a new or different situation. Reading part (b) of the question before starting part (a) would have helped some candidates. Many other answers were excellent and showed that candidates could apply their knowledge to this question. Most candidates focused appropriately on corrective strategies rather than preventive strategies.
- (b) Behaviour modification strategies are based on behaviourist learning theory, and many candidates responded with some excellent answers. Strong responses made the correct distinction between positive reinforcement and positive punishment and negative reinforcement and negative punishment. Weaker responses wrote about Pavlov and his work on dogs, but often failed to write how this related to modifying the behaviour of a disruptive child. Candidates must think about they write and show their understanding rather than learning by rote.

### Question 4

- (a) In order to answer this question, candidates needed to know about both reliability and validity. As mentioned above (Question 2(b)) these are essential terms and candidates should know what they are and be able to apply them in questions such as this one. Any test, whether mathematical or not, should be both reliable and valid. Some candidates did not know the terms (or muddled them) whilst other candidates wrote very good answers considering test-retest and split half reliability and considered a number of different types of validity.
- (b) There was a wide range of answers to this question, with some candidates focusing on testing intelligence (but often did not name an actual test), and others on screening tests (although there was some confusion here). Screening tests are general tests which detect a possible deficit (such as dyslexia) whereas specific assessment tests test a deficit much more precisely.

## *Psychology and Health*

### Question 5

- (a) Many candidates scored full marks for this question part. Although some candidates wrote no more than a brief statement for limited credit, all other candidates wrote a basic statement and expanded on it sufficiently to score full marks. Most answers were 'adherence is the extent to which people carry out the instructions given to them by a medical practitioner', or similar.
- (b) Many candidates also scored full marks for this question part. The studies by Bulpitt on rational adherence and Johnson and Bytheway on customising treatment were most commonly mentioned with many candidates describing these studies in detail. Candidates are reminded that this question part only carries four marks in total, and so long descriptions of studies are not required. A few candidates gave anecdotal answers, e.g. 'people don't take medicine because it tastes nasty', for limited credit. While this is a possible reason, psychological knowledge and evidence need to be demonstrated for full marks.

### Question 6

- (a) Many candidates wrote superb answers, with most of these scoring the maximum mark. Answers covered a wide range of appropriate aspects taken directly from the syllabus which included types and theories of pain, ways in which pain can be measured and ways in which pain can be managed. There were very few anecdotal answers and the overall quality of answers was high.
- (b) Marks awarded to this question part covered the entire mark range, with some answers being very strong with the range and depth of argument, and others not including any evaluation. The named issue for this question was 'how physiological and psychological factors interact'. This should have allowed candidates to bring in a discussion about the gate control theory, which many of the top answers did, but there were those candidates who were unable to discuss the interaction and so either described physiological or psychological aspects.

### Question 7

- (a) This question asked about the use of an interview and many candidates began their answers with ‘I would design an experiment...’ and an interview was either not considered at all, or it covered superficially. Candidates must answer the question set using the method that is stated. Answers not doing this, or using a different method, will score limited or no marks. To conduct an unstructured interview would be appropriate but a face-to-face structured interview would be better and would lead to candidates to suggest what questions could be asked of the practitioner and so extending the answer. Candidates could also consider ways in which the data gathered could be analysed.
- (b) Nearly all candidates chose to describe the Robinson and West study named on the syllabus. The marks achieved reflected the detail and the accuracy shown by candidates when describing the study. As most candidates stated, more information was disclosed to a computer rather than a practitioner, because of the sensitive nature of the medical problem.

### Question 8

- (a) This question had no method stated and so candidates had a free choice of method. Some candidates suggested giving out a questionnaire, others suggested using physiological measures in different situations. Whilst these suggestions are fine, more marks could be gained if the chosen method is described in more detail. Often candidates suggested more than one method. One method in detail, to show in-depth methodological knowledge, is more likely to result in a strong answer.
- (b) This question required a description of the daily hassles explanation of stress. A number of candidates did not know what this meant whilst others gave anecdotal answers. Strong answers described some of the basics of the explanation and the strongest answers distinguished between hassles and uplifts, stated the top five hassles and top five uplifts, and some gave data from the study to support their answers.

## *Psychology and Environment*

### Question 9

- (a) There were some excellent answers written in response to this question. Many candidates appeared to have learned the definition of a cognitive map by Kitchin (1994) word for word. Whilst this approach enables candidates to score full marks, questions state ‘in your own words’ and candidates can score full marks for showing their understanding and explaining what a term or concept means in their own words, without learning direct quotations or definitions.
- (b) There were many superb answers written in response to this question. Like **Question 5(b)**, many candidates wrote answers that were too long. A suggested six lines of average size writing is sufficient to score full marks. There was a wide range of different studies, with the studies by Walcott, Capaldi, Jacobs and Linman, and Tolman being described.

### Question 10

- (a) There were some excellent answers on the topic area of noise and many answers easily achieved full marks for this question. Whilst such answers were strong with the range, depth, and detail included, many candidates wrote far too much for the available eight marks. Candidates should spend time thinking about how to organise an answer, choosing what to include, and showing understanding, rather than trying to include everything they recall.
- (b) The named issue here was ‘the reductionist nature of some studies’ and Examiners took the widest possible definition of reductionism to give candidates the benefit of doubt. Some candidates did not know what reductionism was and gained limited marks. Candidates could consider the reductionist nature of laboratory experiments for example, discussing the use of many controls and the lack of ecological validity. In addition to this issue, strong answers included many others and such answers were characterised by the way in which their answers were organised by different evaluation issues.

### Question 11

- (a) Candidates had to suggest a field experiment to investigate a cognitive coping strategy to reduce the negative effects of crowding. Whilst most candidates knew the basics of an experiment, most emphasised the procedure (i.e. what participants would do) rather than include other equally important aspects of an experiment. For example, very few candidates described the independent and dependent variables, and very few described the experimental design they would apply (repeated or independent measures). Most candidates mentioned they would select a sample of participants but often described the sample rather than the sampling technique. Attention to these aspects would improve marks.
- (b) Answers covered the whole mark range and there were some excellent answers describing the work of Langer and Saegert and Karlin et al. Some candidates wrote too much for the number of marks allocated here, three marks per study.

### Question 12

- (a) This was not a complex question, but one that would show whether candidates could apply what they had learned. Many candidates suggested creating a situation in which people would think the emergency was real. This is never done by psychologists because not only is it very unethical, people could be killed in the panic that may result. Other candidates suggested using virtual reality simulations both in and outside a laboratory. Some candidates suggested using a laboratory experiment which had psychological realism and then generalising from it.
- (b) Candidates had to describe one natural disaster and one technological catastrophe. Some candidates identified an appropriate event, but a lack of detail meant they often scored just one mark. As three marks were allocated to each event there needed to be reasonable detail describing the event. Another weakness evident in answers was that rather than clearly distinguishing between technological and natural events, some candidates described two of the same type, or described events that were ambiguous. Often the events described were known worldwide many appearing in the International news and in psychological literature.

## **Psychology and Abnormality**

### **Question 13**

- (a) A number of candidates explained the genetic explanation of schizophrenia, and in doing so scored two marks, but then they went on to give an example to support their explanation. Whilst this is good examination technique, this question part is allocated only two marks and in this instance it was not a good strategy. Candidates doing this then realised that what they had done in part (a) was now required in part (b). If these candidates had read both parts of the question before writing, then instead of repeating themselves in part (b) they could work more efficiently, saving time and still achieving full marks.
- (b) Candidates here sometimes repeated what was written in part (a), but in general most answers were very good and candidates answered the question appropriately. Evidence featuring twins was prominent with the Gottesman and Shields' research being the most commonly quoted. A few candidates wrote about the dopamine hypothesis, failing to distinguish between genetic and biochemical explanations.

### **Question 14**

- (a) A significant majority of candidates prepared the topic of obsessions and compulsions extremely well, there were some very good and detailed answers that scored maximum marks. Descriptions often included a wide range of different aspects from the syllabus such as definitions, measures, explanations and treatments. The case study of 'Charles' was often quoted, which was an advantage because it provided candidates with a measure which could be contrasted with a psychometric test (e.g. the Maudsley inventory) in part (b).
- (b) A number of candidates only considered the named issue of 'biochemical (drug) treatments' and so gained limited marks. A number of other candidates assumed 'biochemical (drug) treatments' meant *all* treatments (such as cognitive-behaviour and psychoanalytic therapy) and they too gained limited marks. The strongest answers considered the named issue and in addition issues such as consideration of the ways in which OCD can be measured and that of competing theoretical explanations.

### **Question 15**

- (a) There were two parts to this question: a description of a type one and a type two error; and provision of an appropriate example. Most candidates could describe each type of error correctly, although a few candidates confused the two. The example was often a description of the Rosenhan study and for most this was a logical progression from the description of the errors, with many candidates scoring full marks.
- (b) This was an 'open-method' choice question allowing candidates to apply whatever method they thought appropriate to investigate practitioner decision-making. Many candidates did not focus on the specific question. Instead candidates wrote about a patient consulting with a practitioner (or several) and checking the reliability of the diagnosis, or alternatively having a large sample consult with the same practitioner to check the reliability of the diagnosis. These approaches focused on the diagnosis, the outcome, rather than the decision-making process of how a practitioner arrived at a decision. To do this a questionnaire or interview would have been apposite.

### **Question 16**

- (a) There were two common reasons for limited marks for responses to this question: the suggested investigation did not use a questionnaire and the suggested investigation was not about learned helplessness/attributional style. Despite this, there were some excellent answers that designed a study using a questionnaire and these candidates were able to include examples to demonstrate clear knowledge of learned helplessness.
- (b) There were some excellent answers here with candidates outlining early work by Seligman (i.e. his studies on dogs) before moving on to look at how his later work explained depression in humans. Weaker responses could describe learned helplessness in dogs but who could not make the link to how this applied either to humans or to depression.

## **Psychology and Organisations**

### **Question 17**

- (a) Although most candidates scored some credit by responding simply that 'goal setting theory helps people to set goals', significant numbers were unable to extend beyond this to score full marks. In order to elaborate, candidates could have referred to goals being specific, clear and achievable.
- (b) Very few candidates scored full marks. Candidates did not demonstrate familiarity with Latham and Locke's goal-setting theory. Latham and Locke believe that workers can be motivated through the setting of goals, but this is only effective when the goals adhere to the five principles of: clarity, challenge, commitment, effectiveness, and achievability.

### **Question 18**

- (a) Weaker responses anecdotal, about what candidates thought was satisfaction at work. Stronger answers covered either job design or ways in which satisfaction could be measured but rarely included both. There were very strong responses from candidates who had clearly covered all relevant aspects from the syllabus including 'attitudes to work' covering absenteeism and sabotage, for example.
- (b) Some candidates could not demonstrate any knowledge of psychometric tests, and so although other issues were sometimes included in the answer, not addressing the named issue meant that marks gained were limited. The use of psychometric tests is common in research on organisations, so candidates need to be competent in discussing its strengths and weaknesses. There were strong answers from candidates who were able to demonstrate knowledge about psychometric tests and about other appropriate issues.

### **Question 19**

- (a) There was a range of different suggestions proposed by candidates in response to this question. Some suggested conducting an interview with the people operating the machinery whilst others suggested using CCTV recordings. These answers often lacked methodological knowledge and answers tended to be anecdotal. Candidates should always emphasise methodological knowledge when making suggestions.
- (b) Answers describing types of decision-making errors should have focused on the work of Riggio who outlines four types of error: of omission, of commission, of sequencing and of timing. However, a number of candidates looked instead at the *causes* of errors such as lack of sleep or working a particular shift pattern. A few candidates even looked at the types of personalities (e.g. extroverts) who are said to be more likely to have accidents. Looking at causes of accidents had little to do with decision-making, operator-machine systems or working in an organisation. Candidates who did base their answer on Riggio's types often scored full marks.

### **Question 20**

- (a) This question resulted in quite a large number of answers scoring both maximum and very high marks because of the accuracy of descriptions and the detail present in answers. A few candidates described non-behavioural theories such as 'great-person' and charismatic leaders.
- (b) Candidates needed to conduct an observation of the types of behaviour shown by a leader to gain marks for this question. Weak answers were often little more than 'I would use CCTV' without any comment about what would be observed or how it would be analysed. In stronger answers, some candidates suggested types of behaviour that could be observed, wrote about a tally chart (response categories) and suggested event sampling. There was often inclusion of inter-rater reliability showing that these candidates could demonstrate knowledge about the observational method.

# PSYCHOLOGY

---

Paper 9698/32  
Specialist Choices

## Key Messages

- Candidates should provide answers that equate to mark allocation, so an answer worth two marks should be short and an answer worth eight 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 four 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 just from what has been learned for Paper 3), to 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 four marks, the allocation of marks is 2 + 2, whereas if a description of one study is required for four marks, then the same amount in total should be written as for the 2 + 2 format. Sometimes candidates wrote far too much for two marks, particularly evident in **Question 5(b)**.
- There were some excellent answers for question part **(b)** that were they were relevant, accurate, detailed and most showed good understanding.

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

The quality of many part **(a)** answers was very good with many candidates using what is included in the syllabus, covering a wider range of sub-topic areas and describing studies well and in detail.

An important message, for question part **(b)**, which is often repeated, is that candidates need to ensure that they know the difference between 'describe' and 'evaluate'. **Section B** question part **(a)** is 'describe' and question part **(b)** is 'evaluate'. 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 (who also gain limited marks);
- those who do not evaluate at all (and score no marks) merely describing more information.

It appeared that some centres had advised candidates to use the same four evaluation issues for every question. This strategy is not recommended, as it meant that candidates were writing about issues that just did not apply to the question. For example, a candidate might write 'pain is high in ecological validity' or 'organisations are ethical' both of which have very little meaning. There are many issues that can be applied to every topic area and candidates are advised to think about and choose issues appropriate to the topic area of the question.

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

One question part invites 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 of what 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 candidates 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 often use the term experiment incorrectly when they are conducting an observation or questionnaire.

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) Most candidates scored full marks for this question part by giving an appropriate explanation of giftedness. Giftedness is when a child performs at (shows evidence of) a much higher level of accomplishment compared to others of the same age/environment.
- (b) Nearly all candidates could identify two types of giftedness, but although some candidates could describe these in sufficient detail to score full marks, others could not. For example, a candidate might write 'a child has exceptional mathematical ability' and write nothing more. Those choosing to write about intelligence as a type of giftedness often gave more than sufficient detail needed for a maximum mark.

##### **Question 2**

- (a) There were superb answers from candidates who included all three perspectives, explained the underlying theory of each briefly, and then went on to give examples from research in support. Many answers were of very high quality. In weaker answers, sometimes there was only one perspective or candidates would describe classical and operant conditioning but not relate this to learning.
- (b) The named evaluation issue here was the usefulness of the different perspectives. In stronger answers, candidates were able to apply what they had written in part (a) to this, discussing the usefulness of programmed learning or teaching circles for example. These candidates considered other issues too, as required for full marks. In weaker answers, some candidates did not evaluate, only considered the named issue, or did not consider the named issue at all, and gained limited marks.

##### **Question 3**

- (a) Some answers were anecdotal and needed to show evidence of the psychology learned. Many other answers were excellent and showed candidates could apply their knowledge to this question. Many candidates focused appropriately on the work of Brophy who outlines features of effective and ineffective praise.

- (b) For behaviourists, both people and animals are motivated through reinforcers and punishments. Many candidates wrote excellent answers describing this. In the strongest answers, candidates made the correct distinction between positive reinforcement and positive punishment, and negative reinforcement and negative punishment. In weaker answers, some candidates wrote about Pavlov and his work on dogs, but needed to consider how this work explained motivation.

#### Question 4

- (a) It is essential that candidates write about the method that is specified in the question, alternative methods do not answer the question set. Those candidates using an observation, as required by the question, often suggested a useful design but often these answers lacked precision and showed a lack of understanding of what is involved in an observation. Being ethical, such as obtaining consent, and acquiring a sample are important, but these should be in addition to rather than replacing details of the actual observation. Candidates need to think about what they will observe, how it can be recorded and how it will be analysed.
- (b) Candidates who had studied this sub-topic area described the work of Bennett or Fontana. Bennett distinguishes between a formal and informal teaching style where the formal style should prevent disruption. Fontana outlined a high initiative teaching style and the nature of this should also prevent disruptive behaviour.

#### *Psychology and Health*

#### Question 5

- (a) Stress can be measured in two main ways: physiologically using a recording device or a sample test, and psychologically by use of a questionnaire. Most candidates understood this and explained the meaning of these two physiological measures. However, many candidates incorrectly thought that a recording device was a questionnaire.
- (b) Except for those writing about questionnaires (e.g. Holmes and Rahe) there were some excellent answers in response to this question. For the recording device the studies by Geer and Maisel (using GSR) and by Goldstein (blood pressure) were common, and for the sample test the studies by Lundberg and by Johansson (analysis of urine) were often very well described.

#### Question 6

- (a) Many candidates wrote superb answers, with most of these scoring full marks. Answers covered a wide range of appropriate aspects from the syllabus that included types and reasons, measures, and how non-adherence could be improved. Some weaker responses instead focused on more peripheral studies, using evidence from the patient-practitioner relationship, for example, and the topics covered were not always relevant.
- (b) Marks awarded to this question part covered the entire mark range, with some answers being very strong with the range and depth of argument, and others not including any evaluation. Strong answers demonstrated knowledge of the named issue of validity and related it appropriately to the topic area. Other candidates confused validity with reliability.

#### Question 7

- (a) Some candidates were able to suggest appropriate experiments, for example, using the UAB pain observation scale, to determine the effectiveness of patient controlled analgesia compared to practitioner administered treatment.
- (b) Candidates are always advised to read questions carefully and this is one instance where many candidates did not. Many described the whole range of techniques for managing or controlling pain, including medical, psychological and 'alternatives' as defined by the syllabus. Other candidates correctly focused on medical techniques, as required by the question. Other techniques could not be credited.



### Question 8

- (a) There was quite a range of different suggestions proposed by candidates in response to this question. Some suggested conducting an interview with the person involved in the accident and others suggested giving the person a questionnaire to assess their personality. Yet others suggested an analysis of shiftwork patterns to determine if that was the cause. Another suggestion was that it might be due to some operator error (one of Riggio's four types of error). Finally others suggested using CCTV recordings, although these answers were often not sufficiently clear as to how CCTV would be used. Credit was awarded for methodological knowledge.
- (b) This question required a description of two accidents that could be attributed to 'Theory A'. Theory A attributes the cause of accidents to an individual person. Most candidates chose to describe 'illusion of invulnerability' quoting the example of the sinking of the Titanic for the first example and 'cognitive overload' experienced by an air traffic controller responsible for two aeroplanes crashing in Zagreb. Some candidates incorrectly wrote about system errors, confusing Theory A with Theory B.

### *Psychology and Environment*

### Question 9

- (a) A number of candidates scored limited credit because they defined noise as 'unwanted sound'. Whilst this is correct, it is not an explanation and more detail was required for two marks, such as an explanation of why the sound is unwanted, or the inclusion of a supporting example.
- (b) At least two factors that make noise annoying were required, plus examples. Kryter (1970) outlines three factors: loudness, unpredictability and a lack of perceived control. Borsky adds an additional four factors, so there were seven for candidates to choose from. Many candidates scored full marks. A few candidates guessed or gave anecdotal answers, e.g. 'noise is annoying when it gets on your nerves', which could not be credited.

### Question 10

- (a) There were some excellent answers on the topic area of density and crowding and many answers achieved full marks. There were very few answers looking at the different topic of crowd behaviour. Strong answers included range, depth, and detail, but many candidates wrote far too much for the available marks. Candidates should spend time thinking about how to organise an answer, choosing what to include and showing understanding, rather than trying to include everything they recall.
- (b) The named issue here was 'the use of experiments to gather data' and so candidates had the opportunity to debate much of the evidence in this area. For example, the Calhoun animal study was conducted in a laboratory whereas the studies by Christian and Dubos were not. Other relevant issues could be the reductionist nature of some studies; the use of physiological data (as used in the study by Lundberg) and the reliability and validity of studies can always be used. Stronger answers included at least three different issues whereas weaker answers included just one issue or provided 'general' evaluation.

### Question 11

- (a) There were some weak responses that did not demonstrate psychological knowledge, and were not relevant to the question, e.g. a description of a candidate's perfect house or comment on interior design, which could not be credited. Strong answers were based on the errors made in the design of the Pruitt-Igoe project, the success of Newman's designs, and even incorporated aspects of successful community design.
- (b) Strong answers were based mainly on the work of Newman who designed Clason Point in New York City and Five Oaks, in Dayton, Ohio. This involved the closing of streets, improving lighting, introducing speed bumps and dividing the area into 'mini-neighbourhoods'. In addition Newman believed that opportunities for surveillance were crucial and what he called zones of territorial influence, encouraging a sense of personal ownership in homeowners.

### Question 12

- (a) There were many very creative suggestions that were interesting to read. However, many answers had flaws: despite the question asking for a field experiment some candidates suggested a laboratory experiment (e.g. use of an MRI scanner) or a different method altogether. Many candidates did not include the essentials of an experiment: independent and dependent variables, controls and the experimental design. More focus on methodology would improve marks.
- (b) There were many excellent answers with the majority of candidates scoring full marks. Many answers were too long for the three marks allocated to each animal study described. There was a wide range of different studies, with the studies by Walcott, Capaldi, Jacobs and Linman, and Tolman being described.

### *Psychology and Abnormality*

### Question 13

- (a) Answers in response to this question on cognitive restructuring covered the whole mark range. It is the process whereby a person replaces negative thoughts with positive thoughts, thereby 'restructuring' their thinking. Some candidates scored limited credit for 'changing thoughts' but most candidates scored full marks for including the change from negative to positive.
- (b) Candidates sometimes repeated here what was written in part (a). Candidates needed to go beyond the basic and refer to, for example, self-blame and ineptness schema, and negative automatic thoughts (NATs) about the self, the world and the future (the negative cognitive triad). Some candidates also mentioned Beck's six-stage process and these candidates usually scored full marks.

### Question 14

- (a) There were many superb answers written in response to this question by students who had been very well prepared. Most candidates covered at least two of the sub-topic bullet points from the syllabus and many covered all three. Many answers had depth, detail, and range in addition to showing organisation and understanding and these answers often achieved full marks. One notable point is that DSM-V, published in 2013, no longer categorises types of schizophrenia. The 2017 syllabus has been modified to account for this, but for now if candidates write about types they will be credited.
- (b) A number of candidates only considered the named issue of 'generalisations' and gained limited marks. Some candidates provided no evaluation at all, simply extending their description from part (a) which could not be credited. Other candidates provided 'general' evaluation mentioning the same issue for each study and repeating the same point. Evaluation by evaluation issue is the most effective way to score high marks.

### Question 15

- (a) Some candidates could not suggest how both biochemical and psychological factors could be investigated. One logical way to do this would be to invite a person with a known impulse control disorder to participate. The person can then imagine performing their impulsive behaviour. A record of some physiological aspect such as cortisol levels by saliva sample using a salivette can be taken, or a sample of blood to determine levels of different hormones. At the same time a self-report questionnaire or an interview could be conducted to determine a participant's thoughts at that time.
- (b) Candidates had to describe how psychological and biochemical factors interact during any impulse control disorder. The process begins with an association of an event with positive feelings, which, after being repeated a number of times results in a need. For example, a kleptomaniac experiences the need (psychological) to steal and both during and after successful escape the positive sensations resulting from dopamine (biochemical) lead the person to want to repeat the behaviour. Whilst many candidates understood how biochemical and psychological factors interact, many candidates did not, often writing about one or the other, but not both.

### Question 16

- (a) Many candidates stated simply 'I would use CCTV' without demonstrating awareness of what would be observed or how it would be analysed. Basic answers like this, along with a lack of methodological knowledge, gained limited marks. Some candidates suggested asking family members to observe a person, again with little demonstration of their understanding of methodological details.
- (b) This question required a description of both the psychodynamic and biochemical explanations of obsessive-compulsive disorder. Whilst there were some superb answers scoring full marks, often achieved by providing a balanced answer with detail of both explanations, other candidates could describe one or the other, or write no more than a sentence on one or the other. Some candidates wrote about little Hans, which was incorrect because Hans had a phobia rather than OCD.

### *Psychology and Organisations*

### Question 17

- (a) It was clear that many candidates knew what this term meant and scored full marks writing that 'job analysis is the systematic study of the task, procedure, tools, duties and responsibilities involved a job' (or words to the same effect). Other candidates appeared to be unfamiliar with the term, guessed at it and often scored no marks.
- (b) Those candidates scoring no marks in part (a) also scored no marks at all in this question part. Other candidates wrote incorrectly about decision-making and the selection of personnel. A few candidates scored partial credit by suggesting the use of observation or questionnaire. Other candidates scored full marks by choosing to describe functional job analysis (FJA) and the positional analysis questionnaire (PAQ), both of which are used to analyse jobs.

### Question 18

- (a) Weak answers to this question were often anecdotal and included a list of common-sense factors related mainly to physical work conditions. Strong answers referred to psychological studies across the range of relevant sub-topics including physical and psychological conditions, temporal conditions and ergonomics.
- (b) There were many answers achieving very high marks by candidates covering a number of relevant issues including ecological validity. There were some candidates who could not demonstrate understanding of the term ecological validity or who were confused by the term. In brief, ecological validity is the extent to which something relates to real life. If studies on organisations are conducted within organisations then they must be true of real life.

### Question 19

- (a) There was quite a range of different suggestions proposed by candidates in response to this question, which required a questionnaire to be used to investigate Maslow's three additional needs. Some candidates did not know what these three additional needs were which restricted their answers quite significantly. Other candidates suggested using a structured questionnaire to ask a range of participants relevant questions that related directly to the new needs of cognitive, aesthetic and transcendent. These candidates applied their knowledge of questionnaires and in addition showed good understanding of Maslow's needs hierarchy. These candidates scored some very high marks.
- (b) All candidates attempting this question scored some marks, the basics of Maslow's needs hierarchy were well known. Some candidates scored limited marks for identifying some of his original needs; other candidates scored more marks for identifying the original needs plus describing them. The top marks were awarded to those candidates who also included the three latest needs.

**Question 20**

- (a) Any method could be used for this question. The most logical method was to conduct an observation and record the style of behaviour shown by a leader. Many candidates did this and a few candidates (scoring the highest marks) suggested types of behaviour that could be observed, mentioned using a tally chart (response categories) and suggested the use of inter-rater reliability. All this showed that these candidates had knowledge about the observational method and also about leadership styles.
- (b) The syllabus specifies leadership style as 'Styles: permissive versus autocratic (e.g. Muczyk and Reimann, 1987)'. However, very few candidates knew about this style and the mark scheme was widened to include contingency theory (Fiedler, 1976); situational leadership (Hersey and Blanchard, 1988), and path-goal theory (House, 1979). To receive credit candidates had to focus on style rather than just write about the theory.

# PSYCHOLOGY

---

Paper 9698/33  
Specialist Choices

## Key Messages

- Candidates should provide answers that equate to mark allocation, so an answer worth two marks should be short and an answer worth eight 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 four sides of paper in length.
- Candidates should read all parts of a question before beginning to respond 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 just from what has been learned for Paper 3), to 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 four marks, the allocation of marks is 2 + 2, whereas if a description of one study is required for four marks, then the same amount in total should be written as for the 2 + 2 format. Sometimes candidates wrote far too much for two marks, particularly evident in **Questions 1(b), 5(b) and 9(b)**.
- There were some excellent answers for question part **(b)** that were relevant, accurate, detailed and most showed good understanding.

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

The quality of many part **(a)** answers was very good with many candidates using what is included in the syllabus, covering a wide range of sub-topic areas and describing studies well and in detail.

An important message, for question part **(b)**, which is often repeated, is that candidates need to ensure that they know the difference between 'describe' and 'evaluate'. **Section B** question part **(a)** is 'describe' and question part **(b)** is 'evaluate'. 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 (who also gain limited marks);
- those who do not evaluate at all (and score no marks) merely describing more information.

It appeared that some centres had advised candidates to use the same four evaluation issues for every question. This strategy is not recommended, as it meant that candidates wrote about issues that did not apply to the question. For example, a candidate might write 'pain is high in ecological validity' or 'organisations are ethical' both of which have very little meaning. There are many issues that can be applied to every topic area and candidates are advised to think about and choose issues appropriate to the topic area of the question.

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

One question part invites a candidate to describe and 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 of what 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 requires candidates 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 often use the term experiment incorrectly when they are conducting an observation or questionnaire.

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) Most candidates scored full marks for this question part. The mark scheme required two different aspects of bullying to be included in the explanation such as actions, reasons or examples and nearly all candidates managed to achieve this. One explanation is that bullying is a distinctive pattern of deliberately harming and/or humiliating another person.
- (b) Nearly all candidates gave two explanations for bullying and nearly all candidates scored some credit for each answer. What was often missing was the detail in the answer needed for full marks. Often a candidate could give an explanation, such as 'children bully to seek attention', but there was no additional detail to explain what this meant or there was no example given in support. The range of explanations provided by candidates was vast and any appropriate explanation received credit.

##### **Question 2**

- (a) There was a significant minority of candidates who only considered learning styles, not writing about teaching styles at all. These candidates gained limited marks because they only answered half of the question. At the top end of the mark range candidates often included the learning style models of Curry and Grasha and the types of teaching styles proposed by Bennett and Fontana. Measures of styles were less commonly included but ways in which learning could be made more effective was often included and details of it were competently described.
- (b) The named evaluation issue here was 'the reliability and validity of measures' and this issue has appeared in other Education questions because it is a very important issue for this option. Some candidates addressed the issue well and wrote impressive answers. Some candidates confused the terms and some candidates did not attempt the issue at all. To confirm: reliability is the consistency of the measurement. Reliability of a questionnaire/test (e.g. ASI or Kolb's styles) can best be tested using test-retest. Validity is whether a test measures what it claims to measure: learning styles should measure learning styles. A reminder that this named issue is just one of the advised minimum three issues an answer should include.

### Question 3

- (a) Many answers were anecdotal and needed to show evidence of the psychology learned. The purpose of this question type is to test application skills, where candidates apply what they have learned to a new or different situation. Reading part (b) of the question before starting part (a) would have helped some candidates. Many other answers were excellent and showed that candidates could apply their knowledge to this question. Most candidates focused appropriately on corrective strategies rather than preventive strategies.
- (b) Behaviour modification strategies are based on behaviourist learning theory, and many candidates responded with some excellent answers. Strong responses made the correct distinction between positive reinforcement and positive punishment and negative reinforcement and negative punishment. Weaker responses wrote about Pavlov and his work on dogs, but often failed to write how this related to modifying the behaviour of a disruptive child. Candidates must think about they write and show their understanding rather than learning by rote.

### Question 4

- (a) In order to answer this question, candidates needed to know about both reliability and validity. As mentioned above (Question 2(b)) these are essential terms and candidates should know what they are and be able to apply them in questions such as this one. Any test, whether mathematical or not, should be both reliable and valid. Some candidates did not know the terms (or muddled them) whilst other candidates wrote very good answers considering test-retest and split half reliability and considered a number of different types of validity.
- (b) There was a wide range of answers to this question, with some candidates focusing on testing intelligence (but often did not name an actual test), and others on screening tests (although there was some confusion here). Screening tests are general tests which detect a possible deficit (such as dyslexia) whereas specific assessment tests test a deficit much more precisely.

## *Psychology and Health*

### Question 5

- (a) Many candidates scored full marks for this question part. Although some candidates wrote no more than a brief statement for limited credit, all other candidates wrote a basic statement and expanded on it sufficiently to score full marks. Most answers were 'adherence is the extent to which people carry out the instructions given to them by a medical practitioner', or similar.
- (b) Many candidates also scored full marks for this question part. The studies by Bulpitt on rational adherence and Johnson and Bytheway on customising treatment were most commonly mentioned with many candidates describing these studies in detail. Candidates are reminded that this question part only carries four marks in total, and so long descriptions of studies are not required. A few candidates gave anecdotal answers, e.g. 'people don't take medicine because it tastes nasty', for limited credit. While this is a possible reason, psychological knowledge and evidence need to be demonstrated for full marks.

### Question 6

- (a) Many candidates wrote superb answers, with most of these scoring the maximum mark. Answers covered a wide range of appropriate aspects taken directly from the syllabus which included types and theories of pain, ways in which pain can be measured and ways in which pain can be managed. There were very few anecdotal answers and the overall quality of answers was high.
- (b) Marks awarded to this question part covered the entire mark range, with some answers being very strong with the range and depth of argument, and others not including any evaluation. The named issue for this question was 'how physiological and psychological factors interact'. This should have allowed candidates to bring in a discussion about the gate control theory, which many of the top answers did, but there were those candidates who were unable to discuss the interaction and so either described physiological or psychological aspects.

### Question 7

- (a) This question asked about the use of an interview and many candidates began their answers with ‘I would design an experiment...’ and an interview was either not considered at all, or it covered superficially. Candidates must answer the question set using the method that is stated. Answers not doing this, or using a different method, will score limited or no marks. To conduct an unstructured interview would be appropriate but a face-to-face structured interview would be better and would lead to candidates to suggest what questions could be asked of the practitioner and so extending the answer. Candidates could also consider ways in which the data gathered could be analysed.
- (b) Nearly all candidates chose to describe the Robinson and West study named on the syllabus. The marks achieved reflected the detail and the accuracy shown by candidates when describing the study. As most candidates stated, more information was disclosed to a computer rather than a practitioner, because of the sensitive nature of the medical problem.

### Question 8

- (a) This question had no method stated and so candidates had a free choice of method. Some candidates suggested giving out a questionnaire, others suggested using physiological measures in different situations. Whilst these suggestions are fine, more marks could be gained if the chosen method is described in more detail. Often candidates suggested more than one method. One method in detail, to show in-depth methodological knowledge, is more likely to result in a strong answer.
- (b) This question required a description of the daily hassles explanation of stress. A number of candidates did not know what this meant whilst others gave anecdotal answers. Strong answers described some of the basics of the explanation and the strongest answers distinguished between hassles and uplifts, stated the top five hassles and top five uplifts, and some gave data from the study to support their answers.

## *Psychology and Environment*

### Question 9

- (a) There were some excellent answers written in response to this question. Many candidates appeared to have learned the definition of a cognitive map by Kitchin (1994) word for word. Whilst this approach enables candidates to score full marks, questions state ‘in your own words’ and candidates can score full marks for showing their understanding and explaining what a term or concept means in their own words, without learning direct quotations or definitions.
- (b) There were many superb answers written in response to this question. Like **Question 5(b)**, many candidates wrote answers that were too long. A suggested six lines of average size writing is sufficient to score full marks. There was a wide range of different studies, with the studies by Walcott, Capaldi, Jacobs and Linman, and Tolman being described.



### Question 10

- (a) There were some excellent answers on the topic area of noise and many answers easily achieved full marks for this question. Whilst such answers were strong with the range, depth, and detail included, many candidates wrote far too much for the available eight marks. Candidates should spend time thinking about how to organise an answer, choosing what to include, and showing understanding, rather than trying to include everything they recall.
- (b) The named issue here was ‘the reductionist nature of some studies’ and Examiners took the widest possible definition of reductionism to give candidates the benefit of doubt. Some candidates did not know what reductionism was and gained limited marks. Candidates could consider the reductionist nature of laboratory experiments for example, discussing the use of many controls and the lack of ecological validity. In addition to this issue, strong answers included many others and such answers were characterised by the way in which their answers were organised by different evaluation issues.

### Question 11

- (a) Candidates had to suggest a field experiment to investigate a cognitive coping strategy to reduce the negative effects of crowding. Whilst most candidates knew the basics of an experiment, most emphasised the procedure (i.e. what participants would do) rather than include other equally important aspects of an experiment. For example, very few candidates described the independent and dependent variables, and very few described the experimental design they would apply (repeated or independent measures). Most candidates mentioned they would select a sample of participants but often described the sample rather than the sampling technique. Attention to these aspects would improve marks.
- (b) Answers covered the whole mark range and there were some excellent answers describing the work of Langer and Saegert and Karlin et al. Some candidates wrote too much for the number of marks allocated here, three marks per study.

### Question 12

- (a) This was not a complex question, but one that would show whether candidates could apply what they had learned. Many candidates suggested creating a situation in which people would think the emergency was real. This is never done by psychologists because not only is it very unethical, people could be killed in the panic that may result. Other candidates suggested using virtual reality simulations both in and outside a laboratory. Some candidates suggested using a laboratory experiment which had psychological realism and then generalising from it.
- (b) Candidates had to describe one natural disaster and one technological catastrophe. Some candidates identified an appropriate event, but a lack of detail meant they often scored just one mark. As three marks were allocated to each event there needed to be reasonable detail describing the event. Another weakness evident in answers was that rather than clearly distinguishing between technological and natural events, some candidates described two of the same type, or described events that were ambiguous. Often the events described were known worldwide many appearing in the International news and in psychological literature.

## **Psychology and Abnormality**

### **Question 13**

- (a) A number of candidates explained the genetic explanation of schizophrenia, and in doing so scored two marks, but then they went on to give an example to support their explanation. Whilst this is good examination technique, this question part is allocated only two marks and in this instance it was not a good strategy. Candidates doing this then realised that what they had done in part (a) was now required in part (b). If these candidates had read both parts of the question before writing, then instead of repeating themselves in part (b) they could work more efficiently, saving time and still achieving full marks.
- (b) Candidates here sometimes repeated what was written in part (a), but in general most answers were very good and candidates answered the question appropriately. Evidence featuring twins was prominent with the Gottesman and Shields' research being the most commonly quoted. A few candidates wrote about the dopamine hypothesis, failing to distinguish between genetic and biochemical explanations.

### **Question 14**

- (a) A significant majority of candidates prepared the topic of obsessions and compulsions extremely well, there were some very good and detailed answers that scored maximum marks. Descriptions often included a wide range of different aspects from the syllabus such as definitions, measures, explanations and treatments. The case study of 'Charles' was often quoted, which was an advantage because it provided candidates with a measure which could be contrasted with a psychometric test (e.g. the Maudsley inventory) in part (b).
- (b) A number of candidates only considered the named issue of 'biochemical (drug) treatments' and so gained limited marks. A number of other candidates assumed 'biochemical (drug) treatments' meant *all* treatments (such as cognitive-behaviour and psychoanalytic therapy) and they too gained limited marks. The strongest answers considered the named issue and in addition issues such as consideration of the ways in which OCD can be measured and that of competing theoretical explanations.

### **Question 15**

- (a) There were two parts to this question: a description of a type one and a type two error; and provision of an appropriate example. Most candidates could describe each type of error correctly, although a few candidates confused the two. The example was often a description of the Rosenhan study and for most this was a logical progression from the description of the errors, with many candidates scoring full marks.
- (b) This was an 'open-method' choice question allowing candidates to apply whatever method they thought appropriate to investigate practitioner decision-making. Many candidates did not focus on the specific question. Instead candidates wrote about a patient consulting with a practitioner (or several) and checking the reliability of the diagnosis, or alternatively having a large sample consult with the same practitioner to check the reliability of the diagnosis. These approaches focused on the diagnosis, the outcome, rather than the decision-making process of how a practitioner arrived at a decision. To do this a questionnaire or interview would have been apposite.

### **Question 16**

- (a) There were two common reasons for limited marks for responses to this question: the suggested investigation did not use a questionnaire and the suggested investigation was not about learned helplessness/attributional style. Despite this, there were some excellent answers that designed a study using a questionnaire and these candidates were able to include examples to demonstrate clear knowledge of learned helplessness.
- (b) There were some excellent answers here with candidates outlining early work by Seligman (i.e. his studies on dogs) before moving on to look at how his later work explained depression in humans. Weaker responses could describe learned helplessness in dogs but who could not make the link to how this applied either to humans or to depression.

## **Psychology and Organisations**

### **Question 17**

- (a) Although most candidates scored some credit by responding simply that 'goal setting theory helps people to set goals', significant numbers were unable to extend beyond this to score full marks. In order to elaborate, candidates could have referred to goals being specific, clear and achievable.
- (b) Very few candidates scored full marks. Candidates did not demonstrate familiarity with Latham and Locke's goal-setting theory. Latham and Locke believe that workers can be motivated through the setting of goals, but this is only effective when the goals adhere to the five principles of: clarity, challenge, commitment, effectiveness, and achievability.

### **Question 18**

- (a) Weaker responses anecdotal, about what candidates thought was satisfaction at work. Stronger answers covered either job design or ways in which satisfaction could be measured but rarely included both. There were very strong responses from candidates who had clearly covered all relevant aspects from the syllabus including 'attitudes to work' covering absenteeism and sabotage, for example.
- (b) Some candidates could not demonstrate any knowledge of psychometric tests, and so although other issues were sometimes included in the answer, not addressing the named issue meant that marks gained were limited. The use of psychometric tests is common in research on organisations, so candidates need to be competent in discussing its strengths and weaknesses. There were strong answers from candidates who were able to demonstrate knowledge about psychometric tests and about other appropriate issues.

### **Question 19**

- (a) There was a range of different suggestions proposed by candidates in response to this question. Some suggested conducting an interview with the people operating the machinery whilst others suggested using CCTV recordings. These answers often lacked methodological knowledge and answers tended to be anecdotal. Candidates should always emphasise methodological knowledge when making suggestions.
- (b) Answers describing types of decision-making errors should have focused on the work of Riggio who outlines four types of error: of omission, of commission, of sequencing and of timing. However, a number of candidates looked instead at the *causes* of errors such as lack of sleep or working a particular shift pattern. A few candidates even looked at the types of personalities (e.g. extroverts) who are said to be more likely to have accidents. Looking at causes of accidents had little to do with decision-making, operator-machine systems or working in an organisation. Candidates who did base their answer on Riggio's types often scored full marks.

### **Question 20**

- (a) This question resulted in quite a large number of answers scoring both maximum and very high marks because of the accuracy of descriptions and the detail present in answers. A few candidates described non-behavioural theories such as 'great-person' and charismatic leaders.
- (b) Candidates needed to conduct an observation of the types of behaviour shown by a leader to gain marks for this question. Weak answers were often little more than 'I would use CCTV' without any comment about what would be observed or how it would be analysed. In stronger answers, some candidates suggested types of behaviour that could be observed, wrote about a tally chart (response categories) and suggested event sampling. There was often inclusion of inter-rater reliability showing that these candidates could demonstrate knowledge about the observational method.