

PSYCHOLOGY

Paper 9698/11
Core Studies 1

Key messages

- Methodology underpins psychology. Candidates need a good grounding in methodological concepts to understand, describe, evaluate, discuss and apply the core studies effectively.
- Candidates need to practise linking ideas, such as controls, designs, strengths and weaknesses, etc. to each of the core studies.
- The central aspects of each core study (its background, aim, procedure, results and conclusions) need to be carefully learned.

General comments

As with all papers, there was a spread of questions on different aspects of the studies, such as background, procedure, results, conclusions and evaluation. In **Section A**, the candidates' knowledge of procedure (**Question 10(a),(b)**, **Question 11(a)**, **Question 12(a)** and **Question 13 (a),(b)**), conclusions (**Question 8(b)**) and evaluation (**Question 10(b)**, **Question 12(b)**) was fairly good. However, some parts of **Section A** of this paper presented particular challenges to some candidates. Many candidates could improve by having a better general understanding of the background of studies (for example in response to **Questions 2(a),(b)** and **4(a),(b)**), whereas a basic understanding of some basic methodological terms that underpin studies was good, (e.g. **Question 5(a)** and **Question 7(a)**). To improve performance further, candidates would benefit from a more effective grasp of methodology in psychology overall, and, specifically, being able to explain how a study illustrates these principles. For example, to be able to improve their answers to questions using details from the study such as about experimental design (**Question 3(a)**), operationalisation (**Question 3(b)**), qualitative data (**Question 5 (a),(b)**), generalisation (**Question 10(b)**) and controls (**Question 14**). On a positive note, many responses showed that candidates could demonstrate the advanced skill of finding similarities and differences (**Question 13(a),(b)**), which requires candidates to be original in their thinking in the face of an unfamiliar question. This is to be encouraged, although linking answers to studies remains problematic.

Overall, better responses were seen to **Question 17** than to **Question 16** in **Section B**. The responses to **Question 16** were less effective because they were not answering the question. Responses typically provided an evaluation of the study not the approach requested. To improve marks, the responses needed to use the study to provide examples for the valuation. However, responses to **Question 17** tended to be more focused, with candidates using their chosen study to some effect to provide examples to illustrate the debate.

Comments on specific questions

Section A

Question 1

- (a) This question part was generally well answered, with a range of features being identified.
- (b) This question part was less well answered than **Question 1(a)**. Although candidates were often able to identify a relevant point from the study in terms of generalisability, few candidates were able to elaborate.

Question 2

- (a) There were some short, concise answers, such as 'Retroactive is when new memories disrupt old memories while proactive is when old memories disrupt new memories.' Many responses were muddled or inaccurate with few clearly expressing the difference between the two types of inhibition.
- (b) An absence of understanding of the key concept of retroactive inhibition meant few responses were creditworthy in this question part. Many responses were simply a repetition of the aim of the study.

Question 3

- (a) Some responses suggested that candidates did not understand the concept of an 'experimental design', so were unable to answer the question. Few incorrectly described a repeated measures design in this question part. Where the response correctly described an independent groups design, the link to the study was not always correct.
- (b) There was a range of possible answers to this question, but it was not well answered. Responses often identified a measure of a dependent variable without describing how it was measured so gained limited credit. Others included only the measure of a dependent variable, without indicating the variable being tested, for limited credit.

Question 4

- (a) Many responses gained both marks on this question part.
- (b) In most cases the response suggested some understanding but this was not always clear. For example, many candidates were able to refer to a result for the passive or active kittens but struggled to explain their answer. Very few responses successfully explained the results in relation to both touch *and* sight. There were some full mark responses, such as: 'The results of the experiment showed that with the visual cliff, a test of depth perception, kittens exposed to both senses walked over the shallow end of the cliff showing that they could tell the depths of the areas beneath them. The experimental kittens, who only received the visual cues from the yoked contraption, walked over every area of visual cliff indiscriminately. This showed that to develop depth perception both senses of visual and tactile were needed.'

Question 5

- (a) Many responses earned limited credit for stating that qualitative data was descriptive (or an equivalent term) but very few were able to elaborate for full marks.
- (b) Many candidates recognised it was necessary to collect qualitative data to measure feelings or to explain the why behind a behaviour but few could expand by relating their example to the study for full marks, i.e. many responses were generic.

Question 6

- (a) Many responses which did not earn credit identified the two conditions of the experiment as aspects of clothing. Other responses successfully identified two features – most referred to the victim wearing a jacket and no tie. A number of responses wrongly described scruffy clothes, etc. One final fairly common incomplete answer was to say that the victim's clothes were the same in both conditions. Whilst this is true, this is not a *feature* of their clothing.
- (b) Many candidates were unable to write a detailed explanation although some candidates who had not earned credit in **Question 6(a)** were able to gain a mark in **Question 6(b)** for a comment relating to looking like ordinary travellers.

Question 7

- (a) Many brief responses referred to 'consistency' or 'to get the same results' but only a minority elaborated for full marks. Often reliability was confused with validity. The responses needed to 'explain', i.e. to provide some detail, rather than to just 'state'.

- (b) This question produced very few full-mark responses. Although a simple statement that the data were numerical, or 'did not need to be interpreted' could have earned credit, few candidates stated this. Even when adequate definitions were presented in **Question 7(a)**, candidates were typically unable to make any useful comment in **Question 7(b)**, which suggests that they would benefit from practise applying their knowledge.

Question 8

- (a) Many candidates knew that Freud collected data from the father and gained the second mark referring to the fact he met Hans once. A small number made errors, such as that Freud used email.
- (b) Many responses were creditworthy for at least limited credit here, saying that he analysed or interpreted the data.

Question 9

Candidates often confused the results of this study with those of studies 1 and 2. Those candidates who did earn credit typically did so by referring to a preference for attractive faces. Many wrongly said infants preferred faces of their own sex.

Question 10

- (a) Most candidates knew this story involved two boys playing with a ball and many gave a little detail to access full marks. A minority of candidates did not *describe* the story but simply stated that it indicated good or bad motives, which in isolation was not creditworthy. Some responses were unnecessarily long.
- (b) Most candidates were able to recognise that a story about boys would not apply well to or be understood well by girls. The second mark was gained by few, suggesting that candidates would benefit from considering 'why not?', for example if it wouldn't apply well to girls, *why not?* Or it may not be understood well by girls, *why not?* By asking themselves such questions, candidates would have given much better answers.

Question 11

- (a) Many correct responses suggested restrictions in sample such as driving experience but few were able to be specific and accurate and state three years minimum. Some responses suggested that candidates had learned the wrong Maguire study.
- (b) Many responses identified that experiments have an IV or DV. However, even when an appropriate feature of an experiment was identified, there was rarely a link to the Maguire study.

Question 12

- (a) Many candidates gained at least some credit here with some gaining full marks by giving an example.
- (b) In general, identification of a problem was well done, although responses relating this problem to the study were not so common. Many candidates mentioned social desirability bias and gained limited credit.

Question 13

- (a) The majority named the case study method. A few gained full marks by referring to a study of one person.
- (b) This question part produced many excellent responses with a wide range of ideas, both obvious and more creative, including ones that were totally original such as 'Freud worked alone but Thigpen and Cleckley worked together so could check their inter-rater reliability'.

Question 14

There were few full-mark responses to this question. Those candidates who earned marks were often able to say that the test was completed online but were rarely able to *describe* this control. A number of incorrect responses referred to the students being from humanities or science departments.

Finally, responses gaining no marks tended to include factors other than controls, such as simply measures of the DV, e.g. 'all participants completed the same questionnaires'. When considering a question on controls, candidates might benefit from thinking 'Would this be likely to change or fluctuate randomly? And how would this variation affect the results?'

Question 15

- (a) Many candidates were able to refer to feelings and physical appearance. There were also many responses which simply referred to concentrating on the 'inside' (or 'outside') without any psychological content, which could not be credited.
- (b) Common responses simply stated that that one group had BDD and the other did not. Very few responses to this question part were worthy of full marks.

Section B

Question 16

Many responses in this case did not answer the question because they evaluated the study and not the approach. Some candidates wrote in great detail providing everything they knew about the study but were still unlikely to earn marks. Some candidates wrote responses that were limited because they referred to points such as the small sample in Dement and Kleitman's study limiting generalisation, they could have gone on to gain marks by observing that as this is a physiological process the essential biological processes of which are shared, a larger sample was not necessary, i.e. making a comment relating to the approach. When strengths and weaknesses were identified, many candidates could relate these to the study but explanations showing an understanding of the approach were mostly absent.

Question 17

The responses to this question were typically better than in **Question 16** although few were in the top band. Many responses contained anecdotal comments rather than focusing on examples illustrating the explanations using evidence from the chosen study. Many responses included too much focusing on defining the debate rather than on discussing the debate in relation to the study. Responses for the Milgram and Haney et al. studies tended to be better than those from the Bandura et al. study.

PSYCHOLOGY

Paper 9698/12
Core Studies 1

Key messages

- 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 both parts of a question, **(a)** and **(b)**, in **Section A**, before beginning to write an answer to ensure that the answers to both question answer the question set.
- Where a question states 'in this study' candidates must relate what they write to the study in question, i.e. give an example from the study.
- Candidates should look to quote psychological knowledge wherever possible. Anecdotal answers will not 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.
- The writing of some candidates is difficult to read and all candidates are encouraged to write legibly. Candidates are reminded that writing in ink that shows through on the opposite side of the paper also makes what is written difficult to read.
- It is helpful to Examiners if candidates answer questions in the order in which they are presented on the question paper, although **Section B** could be done before **Section A**.

General comments

The following errors are frequently made by candidates and addressing these would increase marks significantly.

Candidates confuse basic 'command terms' such as identify, outline, explain and describe. Often writing one more sentence would make a difference to the mark awarded. An examination is an opportunity for candidates to show in detail what they have learned about psychology.

A common error is not to address 'in this study' in questions. There were many instances of this on this paper and it meant that many candidates scored limited credit because they did not fully answer the question. 'In this study' requires the answer to be related to the study in the question; without doing this the answer could relate to any study when it needs to be explicitly linked.

In **Question 16** and **Question 17** many candidates often describe rather than evaluate. Those who do evaluate often do not evaluate according to the issues identified by the question. On this paper, **Question 17** focused on the cognitive approach and many candidates described one of the studies, identified strengths and weaknesses of the study, but did not consider the cognitive approach, which did not answer the question set.

Comments on specific questions

Question 1

- (a) The aim of the Mann et al. experiment was successfully described in most answers. A typical answer referred to the difference between truth telling and lie telling behaviours or to the study being conducted in a high stakes situation. Responses which included both of these features scored full marks.
- (b) Nearly all responses to this question scored some credit, but only a few answers successfully achieved full marks. The reason for this is that a common exam technique error was made. The question ends with the words 'to investigate this aim' which means that the answer must refer to the aim of this study, which was not always done. A common response would be 'if there is only one participant this cannot be generalised to anyone else', which scores some credit, but does not refer to this specific study.

Question 2

Answers scoring maximum marks often mentioned the four male and four female judges, where five out of eight of them had to agree on the right target word, with no more than two judges choosing the same foil. Many answers in response to this question scored no marks. Some answers described why four foils were created rather than two, and this was done to counteract the ceiling effect. Whilst this is true, the question focuses on the piloting of the foils. Some answers repeated the words of the question, 'the foils were piloted on a group of eight judges', which could not be credited.

Question 3

Many answers in response to this question scored full marks by identifying a piece of apparatus followed by an explanation for its use. A wide range of apparatus were mentioned, such as the shock generator, the chair on which the learner sat, the two rooms to separate the learner and the teacher, the one-way mirror, the pieces of paper, and even the hat in which the paper to draw lots was placed. There were also things mentioned that were not pieces of apparatus such as the experimenter, the learner/stooge and the 'prods' given to the teacher. The question required two things and some responses included only one.

Question 4

- (a) There were many answers receiving full marks which included relevant aspects such as: the use of real police, being arrested at home, being handcuffed, searched, and being fingerprinted. Many answers scored no marks because of the inclusion of detail regarding the prison, such as the uniform, being stripped naked, and various aspects in relation to the guards. The question emphasised 'before arriving at the prison' and so answers that were not focused on the question could not be credited.
- (b) Like **Question 4(a)**, a number of answers included examples of what the prison guards did to the prisoners. Again, examples involving prison guards scored no marks because this was not relevant to the question. What did score credit, for example, was that the suspects were blindfolded when waiting in a cell of the police station; that they had not actually committed any offence; and that the participants have 'signed-up' to take part in a study on prison life before they were arrested.

Question 5

- (a) This question required identification of two guidelines and answers only needed to include a few words to identify ethical guidelines. For example, an answer that included 'informed consent' and 'right to withdraw' scored full marks. Any other appropriate ethical guideline would score marks. Notably the ethical guideline is 'no deception' whereas many answers wrote the guideline as 'deception'. Guidelines are there to be maintained, and the guideline isn't to maintain deception, but to avoid it.

- (b) Most answers referred to guidelines that were broken, such as the victims deceiving participants when they fell over pretending to be ill or drunk. Some answers focused on maintaining a guideline when stating that non-of the participants were identified by name and so maintaining confidentiality. Many answers did not score marks because of poor examination technique, which was not to answer the question set: often the words of the question 'for one of these guidelines' were not followed.

Question 6

Most answers scored full marks when responding to this question. For Experiment 1, answers correctly focused on over-estimating and under-estimating the number of dots on screen. For Experiment 2, answers correctly focused on preference of paintings by the artists Klee and Kandinsky. A small number of answers simply referred to 'preference of artists' with no elaboration and so scored limited credit. A few answers included nothing that could be given credit.

Question 7

- (a) This question required candidates to outline one of the four hypotheses proposed by Bandura et al. Many responses did this perfectly, often by quoting one hypothesis word for word. Many other answers only partially stated a hypothesis and so scored limited credit.
- (b) Some answers gave a finding, but didn't refer to how this finding could be useful. All questions should be read through carefully and all components should be answered in full.

Question 8

- (a) Answers scoring full marks identify the sampling technique (such as self-selecting) and provided an outline of what this technique involved. Many answers merely *identified* a sampling technique, such as 'opportunity sample' but as the question required an *outline*, this type of answer scored limited credit. Some answers incorrectly stated that the sampling technique was a case study.
- (b) The question states 'in this study' so to score full marks, answers had to apply the disadvantage to the study of little Hans. Answers giving a generic disadvantage of either an opportunity sample or a self-selecting sample (either of these answers received credit) scored limited credit. Answers giving a disadvantage of a case study could not be credited.

Question 9

The strongest answers used the term IV and stated one or more of the actual IVs in this study (for example attractive or unattractive faces) and in addition to this, stated the DV which was the fixation time of the infant participants. Marks could also be awarded for the using the term control and by giving an example of a control (such as the occluded glasses for parents). Some answers stated that 'the study was conducted in the laboratory' and whilst this is true it is not a defining feature of an experiment because an experiment could also be conducted in the 'field'. Some answers stated that 'it's an experiment because it has an IV and DV' which is correct, but this brief statement is a partial answer scoring just one mark. This question was allocated four marks, and so answers needed to be relatively longer than those allocated just two marks.

Question 10

- (a) Many answers identified a type of brain scanner to score limited credit (PET or MRI), but often did not describe the technique, as the question required, and so did not score full marks. Some candidates identified two techniques, but the question required just one.
- (b) Answers could focus on either the PET or MRI scan. Most answers described that the scanner produces a structural brain scan all that the scanner detects which brain areas were active. Many answers did not refer to the study itself, and a simple 'when recalling route information' (to relate the scanned to this study) would score full marks.

Question 11

- (a) Many answers were confused, or suggesting something much more complex than was actually required. Any situation other than a laboratory would score marks, because ecological validity would be higher. This could include a classroom, a shopping mall, a cafe, any two situations outside a laboratory would score the two available marks. Using a real person wearing perfume would also score marks because this would increase the ecological validity.
- (b) In this question part responses needed to state *why* the ecological validity would be increased. This could be that real people wear perfume and others could smell this perfume in real life situations. This would be true to real life than seeing pictures and receiving smells from machine done in a laboratory.

Question 12

- (a) There were many correct answers in response to this question, with the statement that 'it is the extent to which a study tests what it claims to test'. However, many answers were not correct, and many answers showed that they did not know the difference between validity and reliability. For example one answer stated 'the term validity means how much study is valid or reliable'. Validity and reliability do have a relationship, but it is much clearer to deal with these completely separately to reduce confusion.
- (b) The correct answer to this question is that the participants, the staff in the mental institutions, were unaware the pseudo-patients were not real patients and so they behaved normally toward them. This makes the study valid because the true behaviour of the participants can be observed and recorded. There were many correct answers that matched this example, however there were a number that didn't, and often made general statements, frequently related to reliability.

Question 13

- (a) This question required identification of any two features of the participant, and responses scored one mark for each feature. Some answers described the participant's problems, for example that she was hearing voices, which was not the focus of the question.
- (b) This question part required identification of two problems Eve experienced that could be found in patients with other mental disorders. Some of these could be related to their actual illness, such as headaches, dizziness, blackouts, hearing voices, distress and memory loss, and some features could have nothing to do with the actual illness such as marital conflicts and unresolved personal conflicts. Some answers included features that Eve did not have, which could not be credited.

Question 14

- (a) Some answers correctly identified the design and in addition were able to describe the reason for this, which was that the IV was male and female participants, and scored full marks. Other answers correctly stated 'independent measures' but without elaboration, and scored limited credit. Other answers incorrectly stated 'repeated measures' which could not be credited.
- (b) Most correct answers referred to the lack of order effects, such as fatigue or practice, because a participant only performs the task once. An alternative correct answer was a possible reduction in demand characteristics because the participant only sees one level of the independent variable.

Question 15

- (a) Marks could be scored by mentioning any two features from 'closed questions', 'scored on a scale', 'a scale from 1-5', 'a scale ranging from strongly disagree to strongly agree'. In addition, marks could be scored by giving an example other question, such as 'I look in the mirror to see how I feel'. Many responses to this question scored zero marks because the answer repeated the question by stating 'it was measured by stating 12 statements about beliefs'.
- (b) There were many incorrect answers in response to this question part, with responses often suggesting an observation.

Question 16

Many answers started with introduction to the study they had chosen. Any description of this type scored no marks because there are no marks allocated for description. Answers should focus on evaluation from the start. Many answers did not address the question of reliability, because many just described the study in full or focused on a combination of reliability and validity when these two are very different things. Answers receiving top marks organised their answers and focused on the chosen study being high in reliability (i) because of a standardised procedure, which should be the same for all participants; (ii) because of experimental controls, controlling extraneous variables, such as a situation. In relation to the study by Schachter and Singer for the observations there were two observers and so inter-rater reliability applied. For the studies by Schachter and Singer and Dement and Kleitman, objective measures were used such as pulse rate and EEG. The most common reason for low reliability was the lack of control over the situation (Loftus and Pickrell conducting interviews in different places) all through the use of subjective data (Dement and Kleitman and dream reports).

Question 17

Like **Question 16**, answers to this question sometimes started with an introduction which does not score any marks. Similarly, answers which did nothing more than describe the chosen study scored no marks. Further, the requirement is not a consideration of any strength or weakness of the chosen study. **Section B** questions like this require the consideration of the strengths and weaknesses of the named issue, in this instance the nature versus nurture debate. Examples supporting this debate should be from the chosen study. Answers which focus specifically on the requirements of the question always score the highest marks. Candidates choosing to base their answers on the study by Held in Hein often showed understanding of which aspects were nature (for example change in pupil size) and which were nurture (for example visually guided movement). A number of answers did not know which aspects were nature and which nurture. Ambiguity also existed in relation to the study by Freud. Some answers believed the phallic stage was learned, but as stages of psychosexual development are said to be present for all children this is something that is inherited.

PSYCHOLOGY

Paper 9698/13
Core Studies 1

Key messages

- 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 both parts of a question, **(a)** and **(b)**, in **Section A**, before beginning to write an answer to ensure that the answers to both question answer the question set.
- Where a question states 'in this study' candidates must relate what they write to the study in question, i.e. give an example from the study.
- Candidates should look to quote psychological knowledge wherever possible. Anecdotal answers will not 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.
- The writing of some candidates is difficult to read and all candidates are encouraged to write legibly. Candidates are reminded that writing in ink that shows through on the opposite side of the paper also makes what is written difficult to read.
- It is helpful to Examiners if candidates answer questions in the order in which they are presented on the question paper, although **Section B** could be done before **Section A**.

General comments

The following errors are frequently made by candidates and addressing these would increase marks significantly.

Candidates confuse basic 'command terms' such as identify, outline, explain and describe. Often writing one more sentence would make a difference to the mark awarded. An examination is an opportunity for candidates to show in detail what they have learned about psychology.

A common error is not to address 'in this study' in questions. There were many instances of this on this paper and it meant that many candidates scored limited credit because they did not fully answer the question. 'In this study' requires the answer to be related to the study in the question; without doing this the answer could relate to any study when it needs to be explicitly linked.

In **Question 16** and **Question 17** many candidates often describe rather than evaluate. Those who do evaluate often do not evaluate according to the issues identified by the question. On this paper, **Question 17** focused on the cognitive approach and many candidates described one of the studies, identified strengths and weaknesses of the study, but did not consider the cognitive approach, which did not answer the question set.

Comments on specific questions

Question 1

- (a) The aim of the Mann et al. experiment was successfully described in most answers. A typical answer referred to the difference between truth telling and lie telling behaviours or to the study being conducted in a high stakes situation. Responses which included both of these features scored full marks.
- (b) Nearly all responses to this question scored some credit, but only a few answers successfully achieved full marks. The reason for this is that a common exam technique error was made. The question ends with the words 'to investigate this aim' which means that the answer must refer to the aim of this study, which was not always done. A common response would be 'if there is only one participant this cannot be generalised to anyone else', which scores some credit, but does not refer to this specific study.

Question 2

Answers scoring maximum marks often mentioned the four male and four female judges, where five out of eight of them had to agree on the right target word, with no more than two judges choosing the same foil. Many answers in response to this question scored no marks. Some answers described why four foils were created rather than two, and this was done to counteract the ceiling effect. Whilst this is true, the question focuses on the piloting of the foils. Some answers repeated the words of the question, 'the foils were piloted on a group of eight judges', which could not be credited.

Question 3

Many answers in response to this question scored full marks by identifying a piece of apparatus followed by an explanation for its use. A wide range of apparatus were mentioned, such as the shock generator, the chair on which the learner sat, the two rooms to separate the learner and the teacher, the one-way mirror, the pieces of paper, and even the hat in which the paper to draw lots was placed. There were also things mentioned that were not pieces of apparatus such as the experimenter, the learner/stooge and the 'prods' given to the teacher. The question required two things and some responses included only one.

Question 4

- (a) There were many answers receiving full marks which included relevant aspects such as: the use of real police, being arrested at home, being handcuffed, searched, and being fingerprinted. Many answers scored no marks because of the inclusion of detail regarding the prison, such as the uniform, being stripped naked, and various aspects in relation to the guards. The question emphasised 'before arriving at the prison' and so answers that were not focused on the question could not be credited.
- (b) Like **Question 4(a)**, a number of answers included examples of what the prison guards did to the prisoners. Again, examples involving prison guards scored no marks because this was not relevant to the question. What did score credit, for example, was that the suspects were blindfolded when waiting in a cell of the police station; that they had not actually committed any offence; and that the participants have 'signed-up' to take part in a study on prison life before they were arrested.

Question 5

- (a) This question required identification of two guidelines and answers only needed to include a few words to identify ethical guidelines. For example, an answer that included 'informed consent' and 'right to withdraw' scored full marks. Any other appropriate ethical guideline would score marks. Notably the ethical guideline is 'no deception' whereas many answers wrote the guideline as 'deception'. Guidelines are there to be maintained, and the guideline isn't to maintain deception, but to avoid it.

- (b) Most answers referred to guidelines that were broken, such as the victims deceiving participants when they fell over pretending to be ill or drunk. Some answers focused on maintaining a guideline when stating that non-of the participants were identified by name and so maintaining confidentiality. Many answers did not score marks because of poor examination technique, which was not to answer the question set: often the words of the question 'for one of these guidelines' were not followed.

Question 6

Most answers scored full marks when responding to this question. For Experiment 1, answers correctly focused on over-estimating and under-estimating the number of dots on screen. For Experiment 2, answers correctly focused on preference of paintings by the artists Klee and Kandinsky. A small number of answers simply referred to 'preference of artists' with no elaboration and so scored limited credit. A few answers included nothing that could be given credit.

Question 7

- (a) This question required candidates to outline one of the four hypotheses proposed by Bandura et al. Many responses did this perfectly, often by quoting one hypothesis word for word. Many other answers only partially stated a hypothesis and so scored limited credit.
- (b) Some answers gave a finding, but didn't refer to how this finding could be useful. All questions should be read through carefully and all components should be answered in full.

Question 8

- (a) Answers scoring full marks identify the sampling technique (such as self-selecting) and provided an outline of what this technique involved. Many answers merely *identified* a sampling technique, such as 'opportunity sample' but as the question required an *outline*, this type of answer scored limited credit. Some answers incorrectly stated that the sampling technique was a case study.
- (b) The question states 'in this study' so to score full marks, answers had to apply the disadvantage to the study of little Hans. Answers giving a generic disadvantage of either an opportunity sample or a self-selecting sample (either of these answers received credit) scored limited credit. Answers giving a disadvantage of a case study could not be credited.

Question 9

The strongest answers used the term IV and stated one or more of the actual IVs in this study (for example attractive or unattractive faces) and in addition to this, stated the DV which was the fixation time of the infant participants. Marks could also be awarded for the using the term control and by giving an example of a control (such as the occluded glasses for parents). Some answers stated that 'the study was conducted in the laboratory' and whilst this is true it is not a defining feature of an experiment because an experiment could also be conducted in the 'field'. Some answers stated that 'it's an experiment because it has an IV and DV' which is correct, but this brief statement is a partial answer scoring just one mark. This question was allocated four marks, and so answers needed to be relatively longer than those allocated just two marks.

Question 10

- (a) Many answers identified a type of brain scanner to score limited credit (PET or MRI), but often did not describe the technique, as the question required, and so did not score full marks. Some candidates identified two techniques, but the question required just one.
- (b) Answers could focus on either the PET or MRI scan. Most answers described that the scanner produces a structural brain scan all that the scanner detects which brain areas were active. Many answers did not refer to the study itself, and a simple 'when recalling route information' (to relate the scanned to this study) would score full marks.

Question 11

- (a) Many answers were confused, or suggesting something much more complex than was actually required. Any situation other than a laboratory would score marks, because ecological validity would be higher. This could include a classroom, a shopping mall, a cafe, any two situations outside a laboratory would score the two available marks. Using a real person wearing perfume would also score marks because this would increase the ecological validity.
- (b) In this question part responses needed to state *why* the ecological validity would be increased. This could be that real people wear perfume and others could smell this perfume in real life situations. This would be true to real life than seeing pictures and receiving smells from machine done in a laboratory.

Question 12

- (a) There were many correct answers in response to this question, with the statement that 'it is the extent to which a study tests what it claims to test'. However, many answers were not correct, and many answers showed that they did not know the difference between validity and reliability. For example one answer stated 'the term validity means how much study is valid or reliable'. Validity and reliability do have a relationship, but it is much clearer to deal with these completely separately to reduce confusion.
- (b) The correct answer to this question is that the participants, the staff in the mental institutions, were unaware the pseudo-patients were not real patients and so they behaved normally toward them. This makes the study valid because the true behaviour of the participants can be observed and recorded. There were many correct answers that matched this example, however there were a number that didn't, and often made general statements, frequently related to reliability.

Question 13

- (a) This question required identification of any two features of the participant, and responses scored one mark for each feature. Some answers described the participant's problems, for example that she was hearing voices, which was not the focus of the question.
- (b) This question part required identification of two problems Eve experienced that could be found in patients with other mental disorders. Some of these could be related to their actual illness, such as headaches, dizziness, blackouts, hearing voices, distress and memory loss, and some features could have nothing to do with the actual illness such as marital conflicts and unresolved personal conflicts. Some answers included features that Eve did not have, which could not be credited.

Question 14

- (a) Some answers correctly identified the design and in addition were able to describe the reason for this, which was that the IV was male and female participants, and scored full marks. Other answers correctly stated 'independent measures' but without elaboration, and scored limited credit. Other answers incorrectly stated 'repeated measures' which could not be credited.
- (b) Most correct answers referred to the lack of order effects, such as fatigue or practice, because a participant only performs the task once. An alternative correct answer was a possible reduction in demand characteristics because the participant only sees one level of the independent variable.

Question 15

- (a) Marks could be scored by mentioning any two features from 'closed questions', 'scored on a scale', 'a scale from 1-5', 'a scale ranging from strongly disagree to strongly agree'. In addition, marks could be scored by giving an example other question, such as 'I look in the mirror to see how I feel'. Many responses to this question scored zero marks because the answer repeated the question by stating 'it was measured by stating 12 statements about beliefs'.
- (b) There were many incorrect answers in response to this question part, with responses often suggesting an observation.

Question 16

Many answers started with introduction to the study they had chosen. Any description of this type scored no marks because there are no marks allocated for description. Answers should focus on evaluation from the start. Many answers did not address the question of reliability, because many just described the study in full or focused on a combination of reliability and validity when these two are very different things. Answers receiving top marks organised their answers and focused on the chosen study being high in reliability (i) because of a standardised procedure, which should be the same for all participants; (ii) because of experimental controls, controlling extraneous variables, such as a situation. In relation to the study by Schachter and Singer for the observations there were two observers and so inter-rater reliability applied. For the studies by Schachter and Singer and Dement and Kleitman, objective measures were used such as pulse rate and EEG. The most common reason for low reliability was the lack of control over the situation (Loftus and Pickrell conducting interviews in different places) all through the use of subjective data (Dement and Kleitman and dream reports).

Question 17

Like **Question 16**, answers to this question sometimes started with an introduction which does not score any marks. Similarly, answers which did nothing more than describe the chosen study scored no marks. Further, the requirement is not a consideration of any strength or weakness of the chosen study. **Section B** questions like this require the consideration of the strengths and weaknesses of the named issue, in this instance the nature versus nurture debate. Examples supporting this debate should be from the chosen study. Answers which focus specifically on the requirements of the question always score the highest marks. Candidates choosing to base their answers on the study by Held in Hein often showed understanding of which aspects were nature (for example change in pupil size) and which were nurture (for example visually guided movement). A number of answers did not know which aspects were nature and which nurture. Ambiguity also existed in relation to the study by Freud. Some answers believed the phallic stage was learned, but as stages of psychosexual development are said to be present for all children this is something that is inherited.

PSYCHOLOGY

Paper 9698/21
Core Studies 2

Key messages

Section A

Question 1

Candidates should be aware of the requirements of each question in the exam. For example, if asked to describe they should not include evaluative comments. A few candidates found it difficult to focus on describing the features of the laboratory experiment method and instead gave a number of evaluation points in part **(a)**. Candidates should suggest a simple alternative to the original study in part **(b)** and give clear details of the procedure followed, ensuring that they use the method and the sample described in the question. 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 all candidates practise writing these types of questions in advance. Some did not structure their responses appropriately and could not achieve full marks. For example, if the question asks for strengths and weaknesses then four points must be made (two strengths and two weaknesses). In addition, candidates need to address evaluation points to the approach/issue named in the question. Candidates must refer to the named study in their responses to achieve higher marks although the vast majority of candidates did refer to the Milgram study.

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. In addition, candidates must discuss three separate points for the **part (c)** of the **Section B** essay in order to achieve full marks. Some candidates only discussed one point using the studies as examples, and gave a very lengthy answer that achieved limited credit.

General comments

Many candidates 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 very small 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.

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 were able to achieve some marks on this question. Most candidates received some credit for identifying that a laboratory experiment looks at the effect of the independent variable on the dependent variable. Many achieved a mark for mentioning that laboratory experiments have some controls or are standardised in some way. A small minority of candidates achieved marks by identifying and/or describing the different designs used in experiments.
- (b) The majority of candidates achieved in the 5+ band with a few achieving full marks. The studies were set in schools and the candidates came up with a number of very interesting ideas for testing lying in children. Responses sometimes did not clearly identify the IV and DV of their study and achieved fewer marks. There were a number of unethical procedures such as suggesting children would be beaten or expelled if they lied.

Some responses evaluated their idea in this question and received no credit for this as this is the correct response to **Question 1(c)**.

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

Many discussed issues about ethics of using children, ecological validity, demand characteristics, validity, and reliability (including inter-rater reliability).

Stronger answers focused on fewer points but then went on to develop them thoroughly in context. Weaker answers limited themselves by only allowing a sentence to address an issue, so that although they may have covered a number of points, none were developed. Some candidates only briefly identified issues and did not refer back to the context of their own study.

Question 2

- (a) Most responses achieved at least some credit for this question by mentioning that the individual explanation involves the individual's personality/disposition; this is what is affecting behaviour.
- (b) Many candidates referred to the disposition of the teacher/participant and the effect this might have had on the result. Better responses added to this comment by referring to the 35% of participants who did not obey and stopped earlier. Some candidates offered a situational explanation of behaviour and referred to the 65% of participants going to 450 volts result to back this up. These responses could not be credited.
- (c) The vast majority of candidates achieved some marks in this section. Many were able to describe at least one strength and/or one weakness of the individual/situational explanation of behaviour. Common strengths included the usefulness of the explanation and the high ecological validity of many studies that recreate realistic situations. Comments on weaknesses included difficulties in determining which explanation is correct and the ethical issues with creating very realistic situations.

A number of responses used issues that were not relevant to the explanation such as qualitative/quantitative data as a weakness or ethics as a strength.

- (d) Some candidate responses were appropriate for this question and identified points about the extent to which the Milgram study has applications to everyday life. Common points included linking the extent the study can be applied to everyday life to ecological validity, sample size, controls used in the study and data collection methods.

A significant number of candidates described the study with little reference to ecological validity and therefore achieved very few marks.

Section B

Question 3

- (a) Many candidates achieved some credit for this question by explaining that validity is how accurately the psychologist measures the behaviour in a study. Some achieved full marks with a clear definition. Some responses stated that validity is a measure of how valid something is which lacked sufficient detail and was therefore not credit-worthy.
- (b) Generally a well-answered question, with some responses achieving high marks for clear and somewhat detailed descriptions of behaviours investigated in the three studies given in the question. Many responses were quite brief and lacked depth. For example, for the Haney, Banks and Zimbardo study, few responses mentioned the shift reports, questionnaires, or interviews. For the Bandura study, the actual way the observation was carried out and the specific behaviours looked for were absent. For the Tajfel study, candidates usually only referred to matrices and without referring to maximum joint profit or in-groups/out-groups.
- (c) Most candidate responses could describe at least one problem that psychologists might have when they attempt to make their study valid. Common issues raised included ethical issues (this sometimes proved problematic because candidates would mention different ethical issues so marks were limited as a result), ecological validity, lack of control of variables, and generalisability. Candidates often linked at least one of these problems to a piece of evidence but it was noticeable that when a candidate did this for the first problem, they would then not link the second and third strength to evidence, and therefore achieved lower marks.

Question 4

- (a) The vast majority of responses achieved some credit by stating that a psychometric test is a measure of the mind and some were able to give an example. The most popular example was IQ test.
- (b) The majority of responses focused on quantitative data collected and very few mentioned the qualitative data collected. There were few omissions with exception of the Thigpen and Cleckley study, in which reference to the EEG tended to be omitted, and in the Billington et al study, where reference to the embedded figures task was often omitted.
- (c) The majority of responses were able to identify one weakness with a psychometric test and often gave evidence as an example. Most candidates found it difficult to give more than one weakness. Common comments on weaknesses included: reductionism, validity (this was often described very well, usually with reference to IQ tests), and lack of detail of the result (reason why participant has given a specific answer).

PSYCHOLOGY

Paper 9698/22
Core Studies 2

Key messages

Section A

Question 1

Candidates should be aware of the requirements of each question in the exam. For example, if asked to describe they should not include evaluative comments. A few candidates found it difficult to focus on describing the features of the laboratory experiment method and instead gave a number of evaluation points in part **(a)**. Candidates should suggest a simple alternative to the original study in part **(b)** and give clear details of the procedure followed, ensuring that they use the method and the sample described in the question. 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 all candidates practise writing these types of questions in advance. Some did not structure their responses appropriately and could not achieve full marks. For example, if the question asks for strengths and weaknesses then four points must be made (two strengths and two weaknesses). In addition, candidates need to address evaluation points to the approach/issue named in the question. Candidates must refer to the named study in their responses to achieve higher marks although the vast majority of candidates did refer to the Loftus et al. study.

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. In addition, candidates must discuss three separate points for the **part (c)** of the **Section B** essay in order to achieve full marks. Some candidates only discussed one point using the studies as examples and gave a very lengthy answer that achieved limited credit.

General comments

Many candidates 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 very small 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.

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 were able to achieve some marks on this question. Most candidates received some credit for identifying correct features of a self-report. Very few gave descriptions of the features and many also gave strengths and weaknesses of the self-report method which did not receive any credit.
- (b) The majority of candidates achieved in the 5+ band with a number achieving full marks. Most understood what was meant by self-report and were able to give examples of the questions they asked to the prisoners/guards. However, some did not give the questions and instead created an experiment which is not creditworthy without the self-report method being used. There were a number of replications of the original Haney, Banks and Zimbardo study as well as unethical studies that were conducted in a simulation of a prison.

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

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

Many discussed issues about bias, validity, ethics, strengths and weaknesses of the data collected and generalisability.

Stronger answers focused on fewer points but really developed them in context. Weaker answers limited themselves by only allowing a sentence to address an issue, so that although they may have covered a number of points, none were developed. Some candidates only briefly identified issues and did not refer back to the context of their own study.

Question 2

- (a) Mainly well answered with some candidates achieving full marks by successfully linking thinking processes to behaviour or give an example (e.g. perception). Those who achieved limited credit did so due to not linking either behaviour or an example to the cognitive approach.
- (b) Many candidates did achieve at least one mark for this question for mentioning that the study was about false memory but many were not able to elaborate on how that related to the cognitive approach.
- (c) The vast majority of candidates did achieve some marks in this section. Many were able to describe at least one strength and/or one weakness of the cognitive approach. Common strengths included the usefulness of the approach, explanations produced by the approach and the controlled nature of the lab experiments often used in the approach. Comments on weaknesses included difficulties in measuring the mind and the lack of ecological validity of the laboratory experiments often used in the approach.

A number of candidates used issues that were not relevant to the approach such as qualitative/quantitative data and ethics as a weakness of the approach. Some appeared to have learned strengths and weaknesses and then attempted to make them fit the approach used in the question.

- (d) Some candidates answered appropriately for this question and identified points about the extent to which the Loftus and Pickrell study is ecologically valid. Common points included linking the extent to which the study can be applied to everyday life, to ecological validity, sample size, controls used in the study and demand characteristics.

A significant number of candidates described the study with little reference to ecological validity and achieved very few marks.

Section B

Question 3

- (a) Many candidates achieved some credit for this question by explaining that the individual differences approach is about people being unique and/or different. Better responses gave an example of an individual differences approach and what they study (e.g. mental health issues).
- (b) Generally a well-answered question, with some candidates achieving full marks for excellent descriptions of behaviours investigated in the three studies given in the question. Some candidates did not mention diagnoses for the Rosenhan study, which would have improved their mark.

Some responses only discussed the behaviour from one study at length, rather than mentioning behaviour observed in all three studies.

- (c) Most candidates could describe two problems of investigating individual differences. Common issues raised included ethics, bias and generalisability from small samples. The candidates did often link at least one of these problems to a piece of evidence but it was noticeable that if the candidate did this for the first problem they would then not link the second and third strength to evidence and therefore achieved lower marks.

Question 4

- (a) The vast majority of candidates were able to explain what is meant by generalisability and link this to either the sample or the ecological validity of the study.
- (b) Some candidates described the reasons why a generalisation could be made from the study (or not). Some candidates gave good responses for this part of the question but often described the procedures of the studies rather than focusing on the generalisations made in each study.
- (c) The majority of candidates could identify two problems of making generalisations from psychology research. Many referred to the problems of ecological validity, representative sample and ethnocentrism and often referred to some evidence to back up the problem. Some candidates wrote very lengthy responses using three studies as examples to back up this one problem.

PSYCHOLOGY

Paper 9698/23
Core Studies 2

Key messages

Section A

Question 1

Candidates should be aware of the requirements of each question in the exam. For example, if asked to describe they should not include evaluative comments. A few candidates found it difficult to focus on describing the features of the laboratory experiment method and instead gave a number of evaluation points in part **(a)**. Candidates should suggest a simple alternative to the original study in part **(b)** and give clear details of the procedure followed, ensuring that they use the method and the sample described in the question. 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 all candidates practise writing these types of questions in advance. Some did not structure their responses appropriately and could not achieve full marks. For example, if the question asks for strengths and weaknesses then four points must be made (two strengths and two weaknesses). In addition, candidates need to address evaluation points to the approach/issue named in the question. Candidates must refer to the named study in their responses to achieve higher marks although the vast majority of candidates did refer to the Loftus et al. study.

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. In addition, candidates must discuss three separate points for the **part (c)** of the **Section B** essay in order to achieve full marks. Some candidates only discussed one point using the studies as examples and gave a very lengthy answer that achieved limited credit.

General comments

Many candidates 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 very small 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.

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 were able to achieve some marks on this question. Most candidates received some credit for identifying correct features of a self-report. Very few gave descriptions of the features and many also gave strengths and weaknesses of the self-report method which did not receive any credit.
- (b) The majority of candidates achieved in the 5+ band with a number achieving full marks. Most understood what was meant by self-report and were able to give examples of the questions they asked to the prisoners/guards. However, some did not give the questions and instead created an experiment which is not creditworthy without the self-report method being used. There were a number of replications of the original Haney, Banks and Zimbardo study as well as unethical studies that were conducted in a simulation of a prison.

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

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

Many discussed issues about bias, validity, ethics, strengths and weaknesses of the data collected and generalisability.

Stronger answers focused on fewer points but really developed them in context. Weaker answers limited themselves by only allowing a sentence to address an issue, so that although they may have covered a number of points, none were developed. Some candidates only briefly identified issues and did not refer back to the context of their own study.

Question 2

- (a) Mainly well answered with some candidates achieving full marks by successfully linking thinking processes to behaviour or give an example (e.g. perception). Those who achieved limited credit did so due to not linking either behaviour or an example to the cognitive approach.
- (b) Many candidates did achieve at least one mark for this question for mentioning that the study was about false memory but many were not able to elaborate on how that related to the cognitive approach.
- (c) The vast majority of candidates did achieve some marks in this section. Many were able to describe at least one strength and/or one weakness of the cognitive approach. Common strengths included the usefulness of the approach, explanations produced by the approach and the controlled nature of the lab experiments often used in the approach. Comments on weaknesses included difficulties in measuring the mind and the lack of ecological validity of the laboratory experiments often used in the approach.

A number of candidates used issues that were not relevant to the approach such as qualitative/quantitative data and ethics as a weakness of the approach. Some appeared to have learned strengths and weaknesses and then attempted to make them fit the approach used in the question.

- (d) Some candidates answered appropriately for this question and identified points about the extent to which the Loftus and Pickrell study is ecologically valid. Common points included linking the extent to which the study can be applied to everyday life, to ecological validity, sample size, controls used in the study and demand characteristics.

A significant number of candidates described the study with little reference to ecological validity and achieved very few marks.

Section B

Question 3

- (a) Many candidates achieved some credit for this question by explaining that the individual differences approach is about people being unique and/or different. Better responses gave an example of an individual differences approach and what they study (e.g. mental health issues).
- (b) Generally a well-answered question, with some candidates achieving full marks for excellent descriptions of behaviours investigated in the three studies given in the question. Some candidates did not mention diagnoses for the Rosenhan study, which would have improved their mark.

Some responses only discussed the behaviour from one study at length, rather than mentioning behaviour observed in all three studies.

- (c) Most candidates could describe two problems of investigating individual differences. Common issues raised included ethics, bias and generalisability from small samples. The candidates did often link at least one of these problems to a piece of evidence but it was noticeable that if the candidate did this for the first problem they would then not link the second and third strength to evidence and therefore achieved lower marks.

Question 4

- (a) The vast majority of candidates were able to explain what is meant by generalisability and link this to either the sample or the ecological validity of the study.
- (b) Some candidates described the reasons why a generalisation could be made from the study (or not). Some candidates gave good responses for this part of the question but often described the procedures of the studies rather than focusing on the generalisations made in each study.
- (c) The majority of candidates could identify two problems of making generalisations from psychology research. Many referred to the problems of ecological validity, representative sample and ethnocentrism and often referred to some evidence to back up the problem. Some candidates wrote very lengthy responses using three studies as examples to back up this one problem.

PSYCHOLOGY

<p>Paper 9698/31 Specialist Choices</p>

Key messages

- Candidates should write 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 note that this is a three-hour examination, 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

Significant numbers of candidates used additional sheets of paper. It is essential that each question is clearly labelled. It is beneficial to Examiners for candidates to arrange additional sheets of answer paper in the correct order and number each sheet, and for candidates to answer questions for each option in the correct order.

Section A (all options):

Candidates should write an amount appropriate to the marks allocated. If a description of two studies is needed for 4 marks, the allocation of marks is 2 + 2, whereas if a description of one study is required for 4 marks, then the same amount in total should be written as for the 2 + 2 format. Sometimes candidates wrote far too much for 2 marks.

Section B (all options):

Answers will receive significantly higher marks if the difference between 'describe' and 'evaluate' is known by candidates. **Section B Question (a)** will always be 'describe' and **Question (b)** will always be 'evaluate'. Evaluation is not simply additional description, but 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, rather than reproduce learning.

Candidates who evaluate can be divided into three types: (i) those who evaluate using a number of evaluation issues in addition to the named issue (and these candidates score the highest marks); (ii) those who focus exclusively on the one named issue and have marks restricted, because one issue is not a range (as required by the mark scheme), and (iii) those candidates who exclude the named issue (and achieve limited marks).

Some centres appear to have instructed candidates to use the same four evaluation issues whatever the question. This strategy is not recommended because it meant that candidates were writing about issues that just did not apply to the question. This was most evident for the Abnormality option when candidates stated that 'phobias are reliable and valid', which is not a meaningful comment. There are many other issues that would be more appropriate and candidates are advised to think carefully and choose issues appropriate to the topic area of the question.

Section C (all options):

In general, answers did not always demonstrate sufficient methodological knowledge. This is evident in the ambiguous use of different methods, and in ambiguous comments about sampling and other aspects such as experimental design. Frequently candidates write 'I will use a random sample of participants' but need to include detail about how that sample will be gathered. Sometimes candidates write that their sample will include 50 males and 50 females and be balanced in terms of age range, etc. (describing the sample) but there will be no details about the sampling technique (i.e. how that sample will be gathered) or how the balance will be achieved).

When a question instructs candidates to use a specific method, then that method must be used. Candidates often start with 'I will conduct an experiment' and write nothing further about the IV or DV or controls or apply a design (repeated measures, for example). These are essential features of an experiment and should be included. Candidates often use the term experiment incorrectly when they are conducting an observation or questionnaire. Candidates are also advised to focus on one method in detail rather than having several superficial sentences about a number of different methods. Candidates should always show their methodological knowledge because application of it scores most marks in this section.

Comments on specific questions

Psychology and Education

Question 1

- (a) Most responses were able to provide an outline of the term 'improving learning effectiveness'. Full marks were awarded for responses which provided some elaboration that went beyond a basic statement, such as by providing an example of one or more study skill techniques.
- (b) Many answers wrote in detail about McCarthy's 4-mat system which includes motivation, concept development, practice and application. Although the question required description of the 4-mat system, many answers described an alternative technique, such as PQRST or SPELT, or answers consisted of an anecdotal technique, which could not be credited.

Question 2

- (a) The strongest answers included a wide range of relevant information related to disruptive behaviour, frequently including information from all three bullet points of the syllabus. A number of answers focused on two bullet points but increased the amount of detail to compensate the lack of range. This is a perfectly acceptable strategy. A small number of answers wrote about disruptive behaviour in their own school and the lack of the relevant psychological knowledge was evident.

- (b) The named issue that had to be included in answers was observations, and most responses did include this issue. However, although many answers provided advantages and disadvantages to evaluate observations, often these were not related to disruptive behaviour, when relating an issue to the topic of the question is essential. Many answers appeared to include three pre-prepared issues, some of which were relevant and some not, which is not good exam technique.

Question 3

- (a) In order to investigate the aim of this study a sample of testosterone needed to be obtained from participants and the level measured on some scale. The investigation would also need to determine whether a participant was autistic or not. Logically a correlation between these two variables could be calculated. Some answers based the investigations on this approach, however many investigations were too vague or proposed investigations which could not gather appropriate data. Spending time thinking through the design of the study before starting to write should be encouraged.
- (b) Although the question required description of just one learning difficulty or disability some answers included two, three or more. Credit could only be given for one of these and the best was always credited. Some answers began by considering dyslexia but then moved to dyspraxia and dyscalculia with the assumption often made that these are all the same thing, when they are not. Answers scoring the highest marks considered just one difficulty or disability and often impressed with the depth of knowledge that was shown.

Question 4

- (a) This question required candidates to describe one explanation of learned helplessness. Many answers focused on the study by Seligman, whose original study investigated learned helplessness in dogs. Alternatively, a number of answers focused on the work by Dweck. In order to achieve full marks, answers needed to be more detailed than when this type of question requires two explanations (three marks each). A small number of candidates wrote about both the above explanations, but given the question, only one could be credited so the answer that would score most marks was credited.
- (b) Candidates could choose any method to investigate the aim of this question although the most logical method to choose would be an experiment. The IV could be males versus females and the DV could be the extent of change in learned helplessness, measured on some scale. The experimental design would be independent as participants would either be male or female. An approach like this was taken by many candidates who often scored very high marks. However, many candidates did not and wrote confusing and sometimes incoherent suggestions that involved a range of different methods. Candidates are advised to choose one method and cover it in detail rather than having one or two sentences on several methods.

Psychology and Health

Question 5

- (a) There were two essential components required in this answer: the promoting of health, and the relating of this to schools. Many responses did this and scored full marks. Some answers only partially answered the question and scored limited credit. Some answers were supported with an appropriate example, commonly the study by Tapper et al. (2003).
- (b) Many responses identified two appropriate methods, typically fear arousal and providing information, and either described each method or gave an example of a study that had been conducted using that method. A number of responses did not mention any appropriate method of health promotion, as the question required, often making suggestions that were not based on psychological knowledge. A small number of responses failed to consider schools and described studies related to worksites or communities, which could not be credited.

Question 6

- (a) Many answers scored full marks because of the range of different aspects included, the amount of detail and the quality of what was written. Most of these answers included types and the reasons for non-adherence, different ways in which adherence can be measured, and various ways in which adherence can be improved. Most answers were based on the studies/topics listed on the syllabus because these are the most relevant. However, many answers brought in information from other topic areas such as the patient-practitioner relationship. Whilst some of these were peripherally relevant, some were not. For example, the study by McKinstry and Wang is not about adherence to medical advice.
- (b) To score full marks in this evaluation section, answers had to consider advantages and disadvantages of issues and not merely identify an issue along with an example. This can be illustrated by reference to the named issue. Candidates would often write 'studying adherence is useful' and 'measuring adherence is useful' without elaboration, which is not sufficient evaluation. Reductionism was often used as an additional issue, but was applied to everything, without any focus. For example, responses would state 'a weakness is that the theory of rational adherence is reductionist' without elaboration or any consideration as to why this might be a weakness or acknowledgement that this is just one of many different explanations for non-adherence.

Question 7

- (a) Answers focusing on the question, and writing about strategies such as attention diversion, non-pain imagery, and cognitive redefinition often scored full marks. A number of answers focused on techniques for managing stress, rather than pain, and could not be credited. Similarly, answers writing about non-cognitive strategies, such as medical or alternative techniques scored no marks.
- (b) The most logical way to determine effectiveness is to conduct an experiment (laboratory or field) whereby two strategies are compared, with each strategy being a condition of an independent variable. Whilst many candidates opted for a design like this, many others did not, some designing studies and which could never investigate the aim. It is advised that time is allocated to thinking through the whole investigation before starting to write the answer. This would make the answer more coherent.

Question 8

- (a) For the first component of the question, answers could focus on either pure physiology: stress hormones such as adrenaline and cortisol, blood pressure, heart rate and all the other physiological changes that occur in a stressful situation. Alternatively, answers could focus on the GAS model by Selye. Some answers focused on both these things. For the second component answers had to focus on the effects of stress on health, writing about stomach ulcers and major effects such as heart attack and stroke. Some answers did this very well, but it was absent from many others. Many answers included no more than a few lines of writing in response to the question and scored very low marks.
- (b) Although there were a number of excellent answers, most attempts at this question scored no more than bottom band marks because of a failure to answer the question set. The question required candidates to conduct a correlational study and if the design did not lead to data that could be correlated then this was a fundamental design error. What was needed, for example, was the gathering of quantitative data of life events on scale and quantitative data from some physiological measure, such as blood pressure, also on a scale. The correlation could then be a positive one predicting that a high life events score would significantly correlate with a high blood pressure score.

Psychology and Environment

Question 9

- (a) There were many excellent answers providing accurate descriptions which often provided an example of a study, such as that by Little, who used the simulation method. Many responses did not describe the simulation method correctly, and scored no marks. Such answers provided general definitions of personal space, or thought that simulation was a type of virtual reality.

- (b) A number of answers scored no marks because the studies described, such as those by Middlemist and by Felipe and Sommer, did not use the simulation method. Other answers described a correct study, by Little, but did not appear to be clear that this study used the simulation method. The simulation study by Little involved the use of dolls placed a distance from each other on a piece of paper. Little measured the distance between the dolls, comparing five different national groups.

Question 10

- (a) Stronger answers included relevant information from all three bullet points of the syllabus. Particularly impressive was the distinction made between theory and evidence shown in many answers, and some answers also distinguished between laboratory studies and real-life situations. Organisation of information like this shows a high level of understanding. There were also answers that listed a few studies that were often accurate, but did not show the same level of organisational skills. A few answers wrote about 'architecture' without demonstrating psychological knowledge. These answers scored very few and often no marks.
- (b) Stronger answers considered three or more issues including the named issue of competing theories. Weaker responses often consisted of a number of assertions without elaboration or understanding. For example, an answer would state that Study X is reductionist and Study Y is also reductionist without considering the strengths and weaknesses of reductionism. Answers like this are identifying rather than evaluating.

Question 11

- (a) This question required the design of a questionnaire. Answers varied significantly according to the quality and accuracy of methodological knowledge. A common error was to identify an aspect of methodology but not to explain it, explanation was needed to improve marks. An example of this is the sample and sampling technique. Some responses for this question were a basic statement, e.g. 'I would get a sample of participants'. Other responses were better, using psychological terminology, when stating for example 'I would use a self-selecting sample'. However, in such cases there was no description about how the self-selecting sample would actually be obtained. Responses achieving most marks would identify the sampling technique, describe how participants would be obtained and also describe some features of the sample of participants.
- (b) This question required description of two studies and most answers did include two, with the studies by Loftus (1972) on evacuation messages and Sattler et al. (2000) on preparedness for an emergency event featuring most frequently. The strongest responses described these studies in similar detail. Some answers had an imbalance between the two and some answers only described one study.

Question 12

- (a) A number of answers identified two errors, but did not provide sufficient description to score more than limited credit. For example, a response might be 'augmentation' and 'Euclidean bias' without elaboration. Answers scoring more marks explained what was meant by a term and often gave an appropriate example, or quoted an appropriate psychological study.
- (b) This question specifically required an experiment to be designed, so the use of other methods scored no marks. An essential feature of any experiment is an independent and dependent variable. Some responses described both of these perfectly; some responses got the IV and DV confused and some responses did not include them. In this instance, the IV could be males versus females and the DV could be the number of sketch map errors. Another essential feature is the control of extraneous variables and some responses described controls in detail, with other responses showing no awareness of controls.

Psychology and Abnormality

Question 13

- (a) Most responses were able to provide at least two features of the case study method and scored full marks. It should be noted that not all case studies are longitudinal; not all case studies automatically gather qualitative (or in-depth) data and not all case studies study people who are unique in some way.
- (b) Most responses to this question described the case of 'Charles' by Rappaport (1989). A few alternative studies were described and these were also credited. A few answers described case studies of disorders other than OCD, which could not be credited.

Question 14

- (a) Stronger answers focused correctly on a number of different models such as the medical/biological, behavioural, psychodynamic and cognitive models. Often answers described the assumptions, applications and treatments of each model and illustrated these with an example of an abnormality. A number of candidates described definitions of abnormality which was also creditable. A number of answers scored very low marks, and some answers no marks at all, because the emphasis of the answer was not on models of abnormality as the question required. These incorrect answers often described a range of different abnormalities such as phobias, OCD, schizophrenia and depression.
- (b) In stronger responses there were many issues debated, such as that between drug treatments being effective but also addictive; and that ECT may be an essential treatment, but it also might have undesirable side-effects. More traditional issues, such as the debate between chemical/biological/medical and psychological approaches, were also considered. In relation to the named issue, some answers failed to go beyond the basic statement of the obvious, for example 'treatments for abnormality are useful' with little or no elaboration. Answers failing to focus on the question set in part (a) had nothing of relevance to evaluate in part (b).

Question 15

- (a) Most answers scored very high marks as two case studies were often described in full and accurate detail. Most answers described the study by Watson (1920) on little Albert and his initial fear of white rats, and the study by Freud (1909) on little Hans. A number of answers did not score marks because they focused on the life of little Hans, the phallic stage and the Oedipus complex rather than on his phobia of horses.
- (b) Every answer scored at least some credit by knowing what was meant by the term systematic desensitisation. However, marks then varied either because knowledge of what this therapy involves was limited or because it wasn't appropriately applied to requirements of the question. An essential feature of systematic desensitisation is the construction of an anxiety hierarchy and then the application of relaxation techniques to remove any anxiety. Many answers failed to mention relaxation techniques at all and so did not access the top mark bands.

Question 16

- (a) A number of candidates designed studies which would not sufficiently investigate the side effects of drug treatments and this lack of coherence in the design prevented them from accessing the top mark bands. For example, many answers simply looked for 'side effects' without showing any awareness of what these side-effects might be. This is where knowledge of drug treatments should be applied in the design of the study. Some designs used a questionnaire to obtain data but the same weakness was evident here because a frequent question to a participant was 'do you have any side-effects?' Other designs used an experiment to compare participants with and without drug treatments and some designs planned to observe side effects in their participants, even for 24 hours a day in some cases.
- (b) Many answers were very detailed and showed excellent understanding when writing about acceptable alternatives which included electroconvulsive therapy (ECT) and psychological treatments such as cognitive restructuring (Beck) and rational emotive therapy (Ellis). A number of answers scored no marks because they focussed on chemical/drug treatments, which the question excluded.

Psychology and Organisations

Question 17

- (a) There were many anecdotal answers in response to this question, although some answers did relate what they wrote to organisations and were awarded credit. One psychological definition is that by Carron (2003) which is ‘the tendency for a group to be in unity whilst working towards a goal’. This unity or cohesion can be based on factors such as social relationships, task relationships, perceived unity and emotion.
- (b) Most answers scored full marks when responding to this question. Detailed answers were often provided which clearly described Tuckman’s five stages of forming, storming, norming, performing, and adjourning. A few answers mention the work of Woodcock and these answers also received credit.

Question 18

- (a) Many answers were of very high quality showing excellent understanding and often very detailed psychological knowledge of a wide range of aspects related to the selection of people for work. Some responses were rather basic, and were unable to demonstrate psychological knowledge, giving common-sense statements about the processes of applying for a job.
- (b) The named issue for this question was interviews, and so candidates could apply methodological knowledge about interviews from any part of the course to this topic. Many responses considered the advantages and disadvantages of interviews as applied to the selection of people for work. Other issues such as the use of psychometric tests, equal opportunities and potential bias in both selection decisions and performance appraisals were often considered. Other responses needed to elaborate beyond the basics.

Question 19

- (a) This question required candidates to describe two questionnaires/rating scales to measure job satisfaction. Although the Job Description Index and the Minnesota Satisfaction Questionnaire are listed on the syllabus, many answers did not mention either of these. Candidates describing these two questionnaires often scored full marks. A number of answers included the designing of their own questionnaire and some answers wrote about the critical incident technique, which could not be credited.
- (b) An interview was used as the primary method by many candidates and some designs were impressive. Some answers used a questionnaire which could not be credited, because the question stated ‘without using a questionnaire’. Many answers did not show sufficient knowledge of interviews to score marks in the top bands, with designs being unclear where the interview would be conducted, for example. Interviews can gather both quantitative and qualitative data if desired but many answers automatically assumed that an interview can only gather qualitative data.

Question 20

- (a) Many responses considered an appropriate range of aspects of observations, often scoring high marks for thoughtful designs. Appropriate aspects included the type (e.g. non-participant), the setting, the use of two observers and the coding (or response) categories. To explain how the study would work, response categories might include a range of different behaviour styles. Some responses were lacking in many (and sometimes all of these) aspects, often with statements limited to ‘I will conduct an observation’.
- (b) There were many excellent answers that correctly described a theory of leadership style clearly and accurately. A number of answers choosing to describe the Muczyk and Reimann theory included a diagram to illustrate how the four styles were determined which was an effective extension to the answer. Many answers briefly distinguished between autocratic and democratic styles with very little elaboration.

PSYCHOLOGY

<p>Paper 9698/32 Specialist Choices</p>

Key messages

- Candidates should write 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 note that this is a three-hour examination, 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

Significant numbers of candidates used additional sheets of paper. It is essential that each question is clearly labelled. It is beneficial to Examiners for candidates to arrange additional sheets of answer paper in the correct order and number each sheet, and for candidates to answer questions for each option in the correct order.

Section A (all options):

Candidates should write an amount appropriate to the marks allocated. If a description of two studies is needed for 4 marks, the allocation of marks is 2 + 2, whereas if a description of one study is required for 4 marks, then the same amount in total should be written as for the 2 + 2 format. Sometimes candidates wrote far too much for 2 marks.

Section B (all options):

Answers will receive significantly higher marks if the difference between 'describe' and 'evaluate' is known by candidates. **Section B Question (a)** will always be 'describe' and **Question (b)** will always be 'evaluate'. Evaluation is not simply additional description, but 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, rather than reproduce learning.

Candidates who evaluate can be divided into three types: (i) those who evaluate using a number of evaluation issues in addition to the named issue (and these candidates score the highest marks); (ii) those who focus exclusively on the one named issue and have marks restricted, because one issue is not a range (as required by the mark scheme), and (iii) those candidates who exclude the named issue (and achieve limited marks).

Some centres appear to have instructed candidates to use the same four evaluation issues whatever the question. This strategy is not recommended because it meant that candidates were writing about issues that just did not apply to the question. This was most evident for the Abnormality option when candidates stated that 'phobias are reliable and valid', which is not a meaningful comment. There are many other issues that would be more appropriate and candidates are advised to think carefully and choose issues appropriate to the topic area of the question.

Section C (all options):

In general, answers did not always demonstrate sufficient methodological knowledge. This is evident in the ambiguous use of different methods, and in ambiguous comments about sampling and other aspects such as experimental design. Frequently candidates write 'I will use a random sample of participants' but need to include detail about how that sample will be gathered. Sometimes candidates write that their sample will include 50 males and 50 females and be balanced in terms of age range, etc. (describing the sample) but there will be no details about the sampling technique (i.e. how that sample will be gathered) or how the balance will be achieved).

When a question instructs candidates to use a specific method, then that method must be used. Candidates often start with 'I will conduct an experiment' and write nothing further about the IV or DV or controls or apply a design (repeated measures, for example). These are essential features of an experiment and should be included. Candidates often use the term experiment incorrectly when they are conducting an observation or questionnaire. Candidates are also advised to focus on one method in detail rather than having several superficial sentences about a number of different methods. Candidates should always show their methodological knowledge because application of it scores most marks in this section.

Comments on specific questions

Psychology and Education

Question 1

- (a) Most responses were able to provide an outline of the term teaching style, but were not always able to provide elaboration or an example for full marks.
- (b) Nearly all responses to this question scored full marks because two teaching styles, based on the work of psychologists, were described. Many answers included the formal and informal teaching styles outlined by Bennett, whilst other answers chose to describe the high initiative and low initiative styles outlined by Fontana. A few answers mentioned all four styles, which was not necessary.

Question 2

- (a) Most responses included a wide range of relevant information related to special educational needs. Some responses focused on just one type of need, such as autism, with others focusing on a range which included autism, ADHD and dyslexia. Some answers also included giftedness which is also a special educational need and so received appropriate credit. Most responses mentioned different strategies for educating children with special needs, such as integration versus segregation.
- (b) The named issue that had to be included in answers was generalisations, and most responses did include this issue. Whilst many responses included other appropriate issues, some responses included issues that just did not apply, or issues that were limited to common sense. A typical illustration of this was the statement 'strategies for educating students with SEN are useful...' which required elaboration in order to be creditable. Many responses used reductionism as an issue, however it was often seen as exclusively negative and as something that automatically applied to everything. A typical comment was 'giftedness is reductionist because not all students are gifted', but more thought was required to improve the overall quality of answers.

Question 3

- (a) Some answers used observation to investigate the effectiveness and others decided to design an experiment. Either method, used appropriately, could score marks. Many responses to this question could not be credited because they did not address the question, which required a focus on a corrective technique, and instead focussed on preventative techniques.
- (b) Many responses correctly focused on the techniques of operant conditioning and many made a useful distinction between techniques that result in a behaviour being rewarded compared to techniques that result in a behaviour being extinguished (e.g. use of positive and negative punishment). Other answers described corrective strategies proposed by others, but needed to describe the theory on which these techniques were based, in order to answer the question set.

Question 4

- (a) Many responses considered an appropriate range of aspects of observations, often scoring high marks for thoughtful designs. Appropriate aspects included the type, (e.g. non-participant), the setting, the use of two observers and the coding (or response) categories. To explain how the study would work, responses categories might include a range of different types of talking. Some responses were lacking in many (and sometimes all of these) aspects, with often statements limited to 'I will conduct an observation'. A few candidates did not focus on an observational study, focusing on another method instead, which could not be credited, as it did not answer the question set.
- (b) Most responses were able to describe two types of bullying. Some of these responses were anecdotal but were still able to score some credit, whilst others were more detailed, showed more understanding, and scored more marks. In addition to verbal and physical bullying, cyber bullying was often considered.

Psychology and Health

Question 5

- (a) Some responses scored credit when correctly stating that 'daily hassles are the small, everyday frustrations that cause stress' and many responses scored full marks by providing some elaboration, such as giving a correct example. A small number of responses rephrased the questions without demonstrating knowledge or understanding, by stating 'a daily hassle is a hassle that happens daily', which could not be credited. A few other responses were incorrect because of a description of everyday occurrences, such as going to work.
- (b) Marks were awarded for the quality of description and correct features which included: 100 adults completing the scale once a month for 10 months; there were 117 hassle items; there were 135 uplifts; examples of uplifts also be credited. A number of responses incorrectly wrote about the Holmes and Rahe life event units, which could not be credited as they did not answer the question set. A number of responses repeated their answer for **Question 5(a)**.

Question 6

- (a) Many answers scored full marks because of the range of different aspects included, the amount of detail and the quality of what was written. Most of these answers included types of and theories of pain, different ways in which pain can be measured, and various techniques to manage and control pain. A few answers confused strategies for managing stress with strategies for managing pain which could not be credited. A few answers focused just on one aspect from the syllabus, such as types and theories, and although and this was often done very well, it restricted the issues that could be applied in **Question 6(b)**.
- (b) Many responses to this question could have improved their approach to evaluation in order to score full marks. It is not sufficient to merely identify an issue along with an example, answers must consider its advantages and disadvantages. For example, some candidates write 'managing pain is useful' without elaboration, which is not sufficient. If 'usefulness' is the evaluation issue, more than a common-sense statement is needed in order to score high marks. Reductionism was often applied to everything, without sufficient thought. For example, responses would state 'the paediatric

pain questionnaire is reductionist because it only applies to children'. As the PPQ is specifically designed for children, reductionism is positive in this case.

Question 7

- (a) Any method could have been used to answer this question, but the popular choice was a questionnaire. An important aspect of any design is the sample and sampling technique. Some responses for this question were limited to 'I would get a sample of participants'. Other responses were better, using psychological terminology, when stating for example 'I would use a random sample'. However, in such cases description was required about how the random sample would be obtained. Responses achieving most marks would identify the sampling technique, describe how it would be obtained and also described some features of the sample of participants.
- (b) Most answers in response to this question scored very high marks. Descriptions were usually focused on the study by Savage and Armstrong, but the study by Byrne and Long was also sometimes described. Some answers had very little detail and only scored minimal marks, whereas others were often extremely detailed showing a very thorough knowledge and understanding of the study. A number of answers scored no marks when describing the study by McKinstry and Wang. This study is not about patient-centred/doctor-centred practitioner styles.

Question 8

- (a) This question required an experiment to be designed, so the use of other methods could not be credited. An essential feature of any experiment is an independent and dependent variable. Some responses described both of these perfectly; some responses confused them and some responses did not include them. Another essential feature is to control extraneous variables and again some responses described controls in detail, with other responses showing no awareness of controls..
- (b) Appropriate answers described the illusion of invulnerability, personality type (such as introverts and extraverts, and the accident prone personality) or other features causing mental impairment that might lead to an error/accident. A number of responses described the study by Barber (1988) on cognitive overload, which could not be credited, because the question excluded cognitive overload.

Psychology and Environment

Question 9

- (a) There are two essential components of a sketch map: the actual drawing or sketching of the map (on paper) and secondly it is a map of what people have in their heads, their cognitive map. Inclusion of these two components would score maximum marks. Many answers failed to refer to the drawing/sketching component and so only scored limited credit.
- (b) In order to score marks, answers had to include the typical features of sketch maps as identified by psychologists such as Lynch. Lynch outlined five features: edges, districts, nodes, landmarks and paths. Identification of two (or more) of these features scored limited credit and description of two (or more) features scored full credit. A number of answers scored no marks because they described incorrect features.

Question 10

- (a) There were many excellent responses to this question with answers often being very detailed and covering a wide range of relevant information. Very good answers covered a wide range of different studies and excellent answers organised these studies into different methodologies. An example of this is knowing that the study by Little is a simulation; that the case study of SM by Kennedy used the stop-distance method and the study by Middlemist et al. is invasion of space.
- (b) Stronger answers considered three or more issues including the named issue of generalisations to different cultures. The study by Little featured prominently in discussions of this issue. Many answers failed to consider evaluation issues in full. Many answers were repetitive, for example stating that Study X is reductionist and Study Y is also reductionist, without considering the strengths and weaknesses of being reductionist. Most candidates viewed reductionism as solely negative when it frequently has many more advantages.

Question 11

- (a) It was necessary for candidates to focus the effect of music on performance. The syllabus states 'performance (e.g. the Mozart effect)' so it would have been appropriate to refer to, for example, the research by Rauscher et al. (1999) which showed that listening to a Mozart piano sonata produced significant short-term enhancement of spatial-temporal reasoning in college students. A number of answers did focus on this study and answered the question set, however many others did not, instead writing about music and consumer behaviour, (e.g. studies by North et al.) and about a study in which classical music decreased blood pressure (Chafin, 2004).
- (b) This question required candidates to design an experiment. In addition to describing the essentials of IV, DV and controls, it is also appropriate to explain the experimental design which can be independent or related. Crucially, all these components should be consistent and coherent throughout the answer. Many responses were, but many responses were confused. Spending time thinking through the design of the study before starting to write should be encouraged.

Question 12

- (a) This question required the interview technique to be used as a main method, and, using both closed and open-ended questions, qualitative and quantitative data could be gathered. Most responses did this and scored high marks. A number of answers used alternative methods and a number of answers only gathered quantitative data. Many responses included very little about the interview technique, failing to mention for example whether it would be conducted face-to-face or by telephone; or whether it would be structured, unstructured or semi-structured.
- (b) Many answers scored high and full marks by describing the study by Lundberg (1976) on commuter passengers on trains in Sweden. Some answers described legitimate studies on crowding in prisons such as those by Paulus et al. Some answers confused *crowding* with *crowds and collective behaviour* and so described inappropriate studies which could not be credited.

Psychology and Abnormality

Question 13

- (a) Answers needed to mention 'learning' (the behavioural assumption that all behaviour is learned) and mention classical conditioning or operant conditioning. A number of responses did not focus on the behavioural approach, and instead focused on medical, psychodynamic or alternative explanations, which could not be credited.
- (b) Many responses focused on the behavioural explanation of phobias, related this to classical conditioning, and often mentioned the case study of little Albert. The question also required explanations based on the behavioural model and so any non-behavioural explanation could not be credited. Some candidates scored limited credit by only describing one abnormality, two were required by the question.

Question 14

- (a) The strongest answers focused correctly on depression and mania, as appears on the syllabus. A number of answers scored very limited credit, and some no credit, because the emphasis of the answer was not on abnormal affect, but instead answers focused on a range of different abnormalities such as OCD, phobias and schizophrenia. A number of answers also considered definitions of abnormality.
- (b) Some answers were limited to very basic statements, for example 'treatments for depression are useful', but elaboration was required. Some answers considered the issue of ethics and often stated that 'ECT is unethical'. It is important to realise that there is a difference between the ethics of conducting psychological studies on participants and the ethics of medical procedures/treatments. Some candidates suggested that ECT and drugs should never be given to anyone because they are 'unethical', without acknowledging that these treatments may be the best and the only available treatment. Answers failing to focus on the question in **Question 14(a)** had nothing of relevance to evaluate in **Question 14(b)**.

Question 15

- (a) A number of candidates designed studies which would not sufficiently investigate the side effects of ECT and this lack of coherence in the design prevented them from accessing the top mark bands. For example, many answers simply looked for 'side effects' without showing any awareness of what these side-effects might be. This is where knowledge of ECT should be applied in the design of the study. Some designs used a questionnaire to obtain data but the same weakness was evident here because a frequent question was 'do you have any side-effects?' Many answers did not recognise that as the patients have schizophrenia, an impaired sense of reality, they may not be able to provide answers to questions.
- (b) Many of these answers were very detailed and showed excellent understanding. Acceptable alternatives to ECT included chemical/drug treatments, cognitive behaviour therapy and token economy, with answers often including the studies by Sensky et al. (2000), and Paul and Lentz (1977). A number of answers focussed on ECT, which was excluded by the question, so could not be credited.

Question 16

- (a) This question required the interview technique to be used as a main method. However, some candidates were unable to demonstrate sufficient knowledge about the interview technique, and needed to mention, for example, whether it would be conducted face-to-face or by telephone; or whether it would be structured, unstructured or semi-structured. There was also a lack of awareness that for effectiveness to be assessed there needed to be a control group who had not had aversion therapy or a group who had received some other treatment. Many answers used questions which was superficial. For example, a commonly asked question was 'do you think the aversion therapy has worked.'
- (b) There were some excellent answers: covert sensitisation was often included as was imaginal desensitisation, although many candidates confused what these two treatments involved. There were some basic answers in which knowledge about addition/impulse control disorder and their treatment was very limited.

Psychology and Organisations

Question 17

- (a) Many responses scored full marks, but others appeared not to be familiar with the term.
- (b) Some responses used the classification proposed by Riggio (1990) where errors could be of omission, commission, of sequence, or of timing. A number of responses focused on accidents in general, failing to relate these to operator-machine systems.

Question 18

- (a) Many answers were of very high quality showing excellent understanding and often very detailed knowledge of a wide range of aspects related to job satisfaction such as job design, various ways in which job satisfaction can be measured, and various attitudes towards work. Some responses were purely anecdotal.
- (b) There were a number of weak answers in response to this question. Many answers considered the named issue of individual differences, but needed to discuss why it is important to study individual differences or consider the implications of individual differences for management of an organisation. Some answers did not evaluate; others identified evaluation issues but did not elaborate, for example, providing statements limited to 'it is useful to know that workers are satisfied at work' and 'it can be generalised that workers are satisfied in their work'.

Question 19

- (a) There were two main weaknesses seen in many answers. Firstly, many answers did not appear to know what task-oriented behaviours were. Secondly, knowledge of observations was often very limited. To achieve the top mark bands, answers should include a range of appropriate aspects of observations such as the type (participant etc.), the setting, the use of two observers and the coding (or response) categories. Many answers assumed that by having two observers, the study was reliable. A test of inter-rater reliability assesses how reliable observations are and the result may show high consistency between observers or it may not. Some responses did not design an observational study, as the question required, and could not be credited.
- (b) Many responses to this question scored full marks when describing two appropriate behavioural theories, most commonly the Ohio State studies (initiating structure and consideration) and the University of Michigan studies (task and relationship-oriented behaviours). Some responses incorrectly described charismatic leaders, transformational leaders and/or the great person theory which did not answer the question, and could not be credited.

Question 20

- (a) Candidates were free to choose any method to investigate the positive effects of group conflict, and a wide range of methods were seen. Many responses were weak because anecdotal explanations of group conflict were investigated rather than explanations based on psychological research. Many other answers were lacking in methodological knowledge, and elaboration was needed to achieve higher marks. Higher marks can be gained for explaining why a methodological aspect has been applied than for simply identifying the aspect.
- (b) Some answers showed clear detailed knowledge and scored high marks. Both positive and negative effects of group conflict were required, but some answers considered only one type of conflict and achieved limited marks. Some answers were entirely anecdotal, and needed to show application of psychological knowledge in order to improve.