

PSYCHOLOGY

Paper 9698/11
Core Studies 1

General comments

As with all papers, there was a spread of questions on different aspects of the studies, such as background, procedure, results and evaluation. In **Section A**, the candidates' knowledge of procedure (**2(a)/(b)**, **3(a)**, **7(b)**, **13(a)/(b)**), results (**3(b)**, **5(a)**, **7(a)**, **8**, **9(a)**), and evaluation (**1(b)**, **4(b)**, **5(b)**, **11(b)**) was good. In general, many candidates could improve by having a better general understanding of the background of studies (for example in responses to **Question 10(a)/(b)**), although in some cases this was good (e.g. **Question 6**). To improve performance still further, candidates would benefit from a more effective grasp of methodology in psychology so that they can see how the study illustrates these principles, for example to be able to improve their answers to **Questions 1(a)**, **4(a)** and **11(a)/(b)**, although the understanding of methodology was good in some areas (e.g. **Questions 12(b)** and **14(b)**) and the understanding of ethical implications was generally good (for example **Question 9(b)**). One very common error was to give findings or conclusions when asked for advantages. This is illustrated by suggestions such as that 'It was an advantage because it showed us that ...'. Such answers are simply a statement of the fact that a useful conclusion could be drawn from the study; this is the purpose of research rather than an advantage of it.

In some cases the ability to apply knowledge was very good indeed (for example in **Question 12(a)**), but in other areas it was not so evident (for example in order to answer **Question 14(b)**).

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

Comments on specific questions

Section A

Question 1

- 1(a)** Many candidates gave simplistic, brief answers which could only earn 1 mark, such as 'They have an IV and a DV'. To earn full marks here, they needed to clarify what each was, e.g. that the IV is manipulated and the DV measured. Candidates also often referred to 'control', although rarely explained its role in the experimental method. A small number of candidates misunderstood what they were being asked to do and gave characteristics of the Mann et al. study, for example, 'they used videos'.
- (b)** Many answers to part **(b)** were not in sufficient detail, so could only earn 1 mark. Those candidates who were successful typically related their idea to lying. Some weaker candidates could show that they knew what a case study was but were not able to explain an advantage.

Question 2

- (a)** A common mistake here was to describe what happened in one or other of the interviews. A number of candidates incorrectly said that the participants were asked to talk to their families about the memory (rather than to *avoid* doing so). Nevertheless, there were also excellent answers, detailing either the need to think about the memory or to avoid talking to others, especially the family, about the study.

- (b) Relatively simple answers suggesting why it was important that they did not talk to their relatives were often able to earn full marks here. Other candidates gained full marks for effectively making the point that there would be potential to enhance the false memory.

Question 3

- (a) This question part was typically well answered, with candidates using all three examples of tests (IQ, eyes test and AQ) as well as the number of participants who were able to correctly identify gender to illustrate their answer.
- (b) This question part was not so well answered. Although many candidates were successful in answering correctly, using comparisons between the two groups, between genders or by referring to the negative correlation, marks were often lost as candidates either referred to “better”/“worse” (which is a value judgement, not a result), believed that group 3 scored higher, or gave the results of the eyes test rather than the AQ test.

Question 4

- (a) Most candidates knew this was a laboratory study but did not always correctly refer to the research method of a laboratory *experiment*. Marks were then sometimes earned for reference to controls. Some candidates mistakenly answered in terms of the experimental design.
- (b) There were many correct answers in this question part. However, common errors were to give answers which referred to findings (see above), or to give two brief advantages rather than one detailed one.

Question 5

- (a) Answers here were often weak. Most correct answers referred to stress, wanting to stop and wanting to know that the learner was OK. This was a very straightforward question although it revealed that many candidates were not familiar with the participants’ actual comments. Instead they guessed at what the participant (the ‘teacher’) may have said or used (incorrect) examples from watching films based on the study. Some candidates left this question blank.
- (b) This question part was answered well. Many candidates knew what self report data was, so were able to explain an advantage successfully.

Question 6

This question was generally well answered although responses were sometimes rather general, simply saying ‘to test helping’. There was also a tendency to list many variables that were tested or questions investigated, rather than to describe two reasons in detail. There was also some confusion over the belief that there were both male and female ‘victims’. Nevertheless, candidates seemed to have taken interest in the Kitty Genovese case, and they were able to report this well, and to discuss the reason why this was important in relation to the Piliavin et al. study.

Question 7

- (a) Most candidates were able to earn at least some credit here, for saying ‘maximum difference’, with many scoring full marks.
- (b) There were some excellent answers to part (b) of this question, with candidates demonstrating a sound understanding of the importance of the random allocation.

Question 8

Answers to this question demonstrated some good understanding of Freud’s ideas. The most common sources of evidence were the giraffe dream and the fantasy of his father becoming the grandfather. There was, however, occasional confusion in some candidates who simply described the Oedipus complex rather than the evidence, or who gave details of more than one source of evidence.

Question 9

- (a) This question was well answered by many candidates, although a common weak response was to simply refer to the preference for attractiveness. In other candidates, incorrect answers suggested guessing as they thought that the babies looked for longer at opposite sex faces.
- (b) Part (b) of this question was generally answered in relation to consent from parents although many other equally valid answers appeared in smaller numbers. There was a small minority of candidates who misunderstood the question and wrote about practical problems.

Question 10

- (a) This question part was not well answered by most candidates. There were, however, some candidates who knew Piaget's theory and could apply it to this question so answered well. It is important that candidates understand the essentials of the background for each core study. There were a significant number of candidates who muddled these studies with Bandura et al. and gave an answer relating to aggression.
- (b) Where candidates demonstrated an understanding of Piaget in part (a), they typically followed this up with a good response in relation to Nelson in part (b), but they were in the minority.

Question 11

- (a) There were few good answers for part (a) of this question. Many answers suggested that the candidate was unaware of the meaning of the term 'experimental design' and few recognised it as a repeated measures design. Even fewer could accurately describe this design. It is important that candidates have an understanding of basic research methodology. Of those candidates who correctly identified repeated measures, many went on to suggest that this meant repeatedly waking participants up to ask them questions. Consequently part (b) was not answered well.

Some candidates showed confusion with the concept of the 'research method' here, with candidates saying that this was an experiment. Some candidates also incorrectly gave conclusions here, rather than describing the design.

- (b) Part (b) of this question was not answered well, largely due to a lack of basic knowledge relating to experimental design. As this is a fundamental aspect of a key research method, it is vital that candidates can recognise and explain it. Some candidates incorrectly referred to conclusions here (as above).

Question 12

- (a) This question was well answered by many candidates. The most common response was volunteer sampling, with suggestions for newspaper or office advertisements. There was clear evidence of candidates utilising their knowledge from other core studies. Another common response was opportunity sampling, although there was some confusion with random sampling, as candidates often described opportunity sampling but called it random sampling.
- (b) In general, when part (a) was answered well, part (b) was also answered well. However, the confusion over random sampling led some candidates to make mistakes here. Nevertheless, when random sampling was correctly described in (a), there were good answers in relation to random sampling in (b).

Question 13

- (a) There were few correct responses here, and it appeared that many candidates did not understand or know the study well enough to answer well. Incorrect answers tended to include 'fresh air'.
- (b) Those candidates who answered correctly in (a) tended to score marks in (b) as well. However, incorrect answers, such as those about 'fresh air', tended to focus on controls rather than the importance of the smells being associated with bodies.

Question 14

- (a) Most candidates were able to give examples of data which could earn marks, although some candidates gave confused responses. Some candidates referred incorrectly to the numbers of apparent pseudo-patients 'identified' in the second study, which was quantitative data, but it needed to be collected by the pseudo-patients themselves to answer the question set.
- (b) Part (b) of this question produced a variety of good answers, with lack of demand characteristics being a common response. Some candidates gave conclusions from the study, a common weak response (see above).

Question 15

- (a) Responses to this question were typically weak, although there were some correct answers which demonstrated a clear comprehension of the question and recall of the study. A common incorrect response was to simply describe the aim of the study itself rather than the purpose of the specific question from the study.
- (b) Those candidates who gave correct answers in part (a) almost always scored full marks in part (b), demonstrating their understanding of both the method and the results of the study.

Section B

Question 16

The most popular choice here was the Haney, Banks and Zimbardo study although there were also good answers based on the Bandura et al. study. The small number of candidates using the Schachter and Singer study tended to produce very weak answers. The quality of responses was varied, with many candidates able to identify basic aspects of application or usefulness such as relevance to prison reforms for the Haney et al. study and parenting or television for the Bandura et al. study. Better essays were able to offer examples of these aspects from their chosen study and the strongest essays expanded on these basic ideas with a range of specific recommendations. Such essays were also more likely to be able to discuss reasons why their chosen study may *not* be very useful. Candidates who included no discussion of an absence of usefulness gained limited marks. At the lower end, these ideas were based on methodology but for better essays additional ideas were explored. Some candidates merely evaluated the study in general, earning marks only by coincidence, or just described the study, earning no marks. Such essays often illustrated that the candidates knew the studies well but did not have the skills to apply that knowledge to the question.

Question 17

This question produced a wider spread of choice across the three studies than **Question 17** and answers again demonstrated very good knowledge of the studies. The Thigpen and Cleckley study produced some good answers but also the weakest ones, for example with candidates often describing hypnosis as a psychometric test. Although answers were more balanced here, with candidates seemingly much clearer about the need to address strengths and weaknesses, there was again a tendency to consider these in relation to the study rather than using the study to illustrate an evaluation of psychometric tests in general.

PSYCHOLOGY

Paper 9698/12
Core Studies 1

Key messages

- Candidates should provide answers that equate to mark allocation, so an answer worth 2 marks should be short and an answer worth 10 marks should be correspondingly longer. **Section B** questions are not short-answer.
- For a **Section A** 2-mark answer that has the command word '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 parts are not the same.
- 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 beneficial 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

Candidates confuse command words such as identify, outline, explain and describe and would benefit from more preparation on examination technique in this area.

Candidates need to ensure they address the 'in this study' part when it is stated in questions. There were seven instances of this on this paper and it meant that many candidates could have scored an additional 7 marks. 'In this study' means that the answer must be related to the study in the question; without doing this the answer could relate to any study.

In **Questions 16** and **17**, candidates are required to evaluate, but many describe. Those who do evaluate often do not evaluate according to the issues identified by the question. On this paper **Question 16** focused on generalisations and instead of answering the question many candidates identified strengths and weaknesses of the study, resulting in limited credit.

Comments on specific questions

Question 1

- (a) Most candidates were able to score at least one mark when answering this question. Those candidates scoring one mark often wrote nothing more than "more words for true stories and less words for false stories" whereas candidates scoring full marks often gave the exact number of words used: 138 for 'true' and 49.9 for 'false'. Candidates were also given credit if the numbers they gave were not exact but were very close to these two figures.

- (b) A small number of candidates did not refer to word count, as the question required, and scored no marks, as did those who did not give an advantage. Some candidates wrote that “it is quantitative data” but did not explain how this was an advantage, whereas those candidates who wrote that for example: “it is quantitative so it is more objective and cannot be affected by the interpretation of the researcher” scored the full 2 marks available. A small number of candidates gave two advantages and each was marked with the best credited.

Question 2

- (a) Most candidates scored full marks in response to this question, because 1 mark for each correct aspect of the sample was allocated. This meant that ‘newborn’ and ‘kittens’ scored candidates full marks. Several other aspects also scored marks, such as ‘10 pairs’, 20, and ‘each pair from a different litter’. Comments about the procedure were not creditable, such as how the kittens were allocated to be ‘active’ or ‘passive’.
- (b) A wide range of reasons were acceptable for this question part. One reason is that kittens have a visual system similar to humans in that they show depth perception; another reason is that kittens are easy for researchers to manage. Also acceptable were comments relating to the kittens being fast learners, being neonate/newly born, or simply that they were mammals. One mark was allocated to a correct but basic answer, such as “they are mammals” whereas answers needed some elaboration for two marks.

Question 3

- (a) The experimental design used in this study was independent groups/between participants/subjects because each participant performed in only one of the levels of the independent variable (IV). In other words, the IV compared those with autism (group 1) with other groups of people without autism (groups 2, 3 and 4). A participant could not be autistic *and* non-autistic. Participants were matched on a number of things, but this was not the main design of the study. Whilst many candidates scored full marks, many did not either because they got the design wrong, or because they did not even know what was meant by the term experimental design.
- (b) This answer in question part (a) had implications for this question part, because if the design was not known then an explanation for its use could not be provided here. As was explained above, a participant could only be autistic or non-autistic and so any other design was inappropriate. For example, a *repeated measures design* could not be used as participants could not be both control and autistic; a *matched pairs design* is hard to use because other than IQ it is not easy to match someone who has autism with someone who has not.

Question 4

- (a) The sampling technique used in this study was a volunteer or self-selected sample. Naming this technique scored 1 mark, with a further mark allocated to describing it. The description could include how the sample was acquired and in this study participants responded through newspaper advertisement. A number of candidates did not know the sampling technique, often stating incorrectly that it was an opportunity or a random sample. Some candidates stated that it was a ‘newspaper sample’, but this is incorrect because this is not a sampling technique; it is just one way in which a researcher can acquire a volunteer/self-selecting sample.

- (b) Most candidates did not address the 'in this study' part of the question and this meant they could not score more than one mark. It is essential that if any question ends with the words 'in this study' then the question must be related to, or an example given from, the study being asked about. If a candidate writes "one advantage is that it may be representative of the wider population" it has not tailored the response to the study. For this question part most candidates scored just one mark because of this.

Question 5

- (a) This question part required candidates to describe one difference between the results from the drunk and ill victims. Those candidates scoring 1 mark often wrote that "more people helped the ill victim than the drunk victim" without elaboration. Those candidates scoring 2 marks provided elaboration for example "ill victims were helped more spontaneously such as 62/65 compared to the drunk who was helped on 19/38 trials". Answers which show more detailed psychological knowledge (i.e. elaboration), such as this latter example, will always score more marks than answers with no elaboration.
- (b) This question part invited candidates to explain why there was a difference in the results for the drunk and ill victims. Those scoring 2 marks usually made a comment about both the ill victim, such as "being ill is not their fault" and in addition a comment about the drunk victim, such as "being drunk is of their choosing". Answers scoring 1 mark often referred to just one of these explanations and some answers that were not creditable said "it is explained by the cost benefit matrix" but then made no further comment about how this applied to the drunk or ill victim.

Question 6

- (a) To score marks for this question, candidates had to identify any two features of laboratory experiments. Many possible points could be credited and these included: IV and DV; control (of variables); IV manipulated; DV measured; artificial environment and that it investigates causal relationships. Some candidates wrote that a feature of an experiment is having a hypothesis, but this was not credited because this feature is not unique to experiments. Candidates often made the assumption that laboratory experiments are automatically low in ecological validity, this is not true.
- (b) Whilst the answer to question part (a) did not need to be contextualised (related to the study in question), this question part did. Those candidates who did not know the features of an experiment in part (a) struggled with this question part. Those who provided a correct answer in (a) did much better, but often struggled to apply the features correctly. Those candidates scoring full marks correctly identified a feature, e.g. the IV and the DV, and then applied these so the IV was the in-group/out-group; and the DV: allocation of points.

Question 7

- (a) There were significant numbers of candidates who did not know the term 'repeated measures design'. Incorrect responses included: "it is when the experiment is repeated by participants", and "it is where a participant repeats the same condition". A repeated measures design is when any one participant (or group) performs in each (or all) of the conditions of the independent variable. This means that a participant does not actually repeat anything.
- (b) The correct answer is that if a participant had seen both a male *and* a female model, or they had been in the no model, aggressive model *and* non-aggressive model conditions then the children would at very least be confused (worth 1 mark) or there might be order effects which would reduce the validity of the whole study. Writing "they would be confused" is correct but shows no psychological knowledge. Writing about order effects or validity (for example) shows good psychological knowledge and scores full marks. Notably the question requires 'in this study' so if there is no reference to the study then 1 mark maximum is awarded. To address the 'in this study' aspect, a candidate could write 'If the children are exposed to the no model, aggressive model *and* non-aggressive model conditions order effects would probably occur and the experiment has no validity as it would not be possible to determine whether the child has copied a model or not'.

Question 8

- (a) There were many different suggestions made by candidates as to what the research method was and the most common suggestions included observation, self report and longitudinal. However, the correct answer was that it was a case study which studies one individual (or one 'unit'). As this case study used a range of different methods (such as those listed above) candidates could have mentioned these and scored marks for elaboration. Note that 'longitudinal' is not a method as such (and neither is a 'snapshot' study).
- (b) This question asked candidates to explain why (i.e.) a case study was the best research method for this study. Those candidates answering (a) correctly had little difficulty in explaining the advantage of this method when studying little Hans.

Question 9

- (a) Variables can be *manipulated* and variables can be *controlled*. In this study the only independent variables (IVs) to be manipulated were Gender (males and females); Age (babies and adults); and Race (Black and White). Inclusion of any two of these three scored full marks. Some candidates gave a range of variables that were *controlled*, which could not be credited.
- (b) The two ways in which infant facial preferences could be acquired are through either innate/inborn/ 'nature' or through learning/'nurture'. Many candidates correctly identified these two ways and scored full marks. Some candidates identified just one of these and scored 1 mark. Some candidates wrote that facial preference is "acquired through forming prototypes" but needed to state how these prototypes are acquired for this to be creditable.

Question 10

Candidates would benefit from a stronger understanding of the distinction between results, findings and conclusions. Although this question asked for conclusions, candidates also received credit if they wrote about findings. In this study two general conclusions were: morality develops (e.g. 7 years olds are better than 3 year olds) and children develop morality quite early (3 year olds could make moral judgments). Many more specific conclusions could be drawn such as: making a moral judgment requires understanding of the evaluative concepts, i.e. good and bad, to be applied; a child must also be able to recognise and interpret the inter-relationship between actions, motives, goals, and outcomes in order to make evaluative judgments; and children as young as 3 years of age can and do use motive information for making moral judgments, but only when this information is explicit and salient. Any other appropriate conclusion received credit. Whilst many candidates scored full marks, others did not and this was for two main reasons: (i) what was written was not a conclusion and sometimes answers included numbers (results) of what children were able to and not able to do; (ii) conclusions were too detailed in that a candidate would write about three or more conclusions, although a maximum 4 marks were available.

Question 11

The most common reason for this study, understood by nearly all candidates, was that it was done to test the two-factor theory of emotion. Many candidates correctly provided elaboration of this theory and the inclusion of a physiological component (whether this be arousal/epinephrine) for an additional mark. Also crucial to the theory is a cognitive component and inclusion of this factor could score an additional mark. Mention of the cognitive labelling of arousal could also be credited. This question was allocated 4 marks, some candidates gave answers that were too brief for this number of marks.

Question 12

- (a) Candidates could describe any control that was used in the study. Most candidates wrote about the same flow rate, the same presentation duration or the same perceived odour intensity. Many candidates also wrote about the brief questionnaire that asked participants about their normal sense of smell and that they had no olfactory dysfunction. Some candidates wrote about more than one control (all marked and the best one answer credited), and some candidates wrote not about controls but the variables (IVs) that were manipulated. 1 mark answers identified the control (e.g. presentation duration) and 2 mark answers gave some elaboration (e.g. that the duration was for 500 milliseconds).

- (b) This question ends with the words ‘in this study’ and so to score both marks this component must be addressed. Many candidates did not relate their response to this study. A common response was “it would confound the study”, scoring limited credit because the answer makes no reference to *this* study. Those candidates writing “failing to control odour intensity might confound the study because any stronger smell would be more obvious and so some participants would notice it more easily” would score full marks.

Question 13

- (a) The most typical examples of qualitative data were queuing early for food, interpreted as oral-acquisitive behaviour; note-taking; engaging in writing behaviour; and walking corridors, perceived as a sign of nervous behaviour. Some responses included quantitative data, which could not be credited.
- (b) Most candidates were able to suggest an appropriate advantage, such as “it gives rich, in-depth information” but many candidates did not relate their response to this study, as required by the question, and received limited credit.

Question 14

- (a) Whereas many candidates answered the question as stated, i.e. what was known at the *start* of the study, many other candidates simply wrote about information that was found out *during* the study. For example, at the start of the study Thigpen and Cleckley knew that she was a 25 year-old woman, was married (with difficulties), had a child, and that she had severe headaches. At the start of the study nothing was known about her shopping trip or her IQ score, or Eve Black for example, these features were discovered during the interviews and tests, and could not be credited.
- (b) This question was answered correctly by most candidates.

Question 15

This question allowed candidates to use a diagram to support their answer. Many candidates provided reasonably accurate diagrams, whereas others tried, did not succeed and crossed out their answers. Others did not provide a diagram. Many features could score marks such as: Figures had a pair of diagrams and a small shape; there were 12 of them; the task was to find a shape (simple/small/black and white); hidden in one of the big/complex/coloured shapes; in pairs of diagrams of which there were 12. Participants could respond by pressing a key in a maximum of 50 seconds before moving on automatically, all of this done online. Marks were allocated for each feature included in the answer.

Question 16

Some candidates gave stronger responses, providing two (and sometimes more) points illustrating generalisations and these points were clearly supported with relevant examples from the named study chosen. These candidates then went on to provide a balanced answer by providing points which restricted generalisations from being made and again showed excellent understanding by using appropriate supporting examples. Answers like this were rare. Weaker responses showed three types of limitations:

- Writing about all three named studies rather than just one study. In such cases all answers were marked and the best one credited, but it often meant that there was much less detail on any one study.
- Candidates who did not answer the question set, instead writing about strengths and weaknesses of the chosen study, perhaps because they did not know what the term ‘generalisation’ meant, and
- Candidates who provided a descriptive answer only.

There is no difference in marks between which one of the three named studies is chosen. However, thinking about which of the three studies would provide more examples for the candidate in relation to the named issue would allow more marks to be scored. Evidence for this is the number of candidates who began an answer, crossed it out, and started again with a different study.

Question 17

This question invited candidates to evaluate the strengths and weaknesses of a named study and this allowed those candidates who do not focus on a specific issue (as in **Question 16**) to score some marks. However, there were still candidates failing to score marks because the question asked for both strengths *and* weaknesses and any answer that provided *only* strengths *or* weaknesses, or a significant imbalance, received limited credit. Some candidates were not able to provide at least two strengths and two weaknesses in their response.

There were also errors when relating the strength or weakness to the named study and candidates are encouraged to think carefully about what they write. For example for the Dement and Kleitman study, some candidates wrote contradictory statements, e.g. “one strength is that sleeping is ecologically valid” and “sleeping is not ecologically valid”, within their answer. It may be the case that the candidate means that sleeping in a laboratory (rather than in one’s bed at home) is low in ecological validity, but responses would be improved if points are clearly made, avoiding ambiguity or the need for interpretation.

PSYCHOLOGY

Paper 9698/13
Core Studies 1

General comments

As with all papers, there was a spread of questions on different aspects of the studies, such as background, procedure, results and evaluation. In **Section A**, the candidates' knowledge of procedure (**2(a)/(b)**, **3(a)**, **7(b)**, **13(a)/(b)**), results (**3(b)**, **5(a)**, **7(a)**, **8**, **9(a)**), and evaluation (**1(b)**, **4(b)**, **5(b)**, **11(b)**) was good. In general, many candidates could improve by having a better general understanding of the background of studies (for example in responses to **Question 10(a)/(b)**), although in some cases this was good (e.g. **Question 6**). To improve performance still further, candidates would benefit from a more effective grasp of methodology in psychology so that they can see how the study illustrates these principles, for example to be able to improve their answers to **Questions 1(a)**, **4(a)** and **11(a)/(b)**, although the understanding of methodology was good in some areas (e.g. **Questions 12(b)** and **14(b)**) and the understanding of ethical implications was generally good (for example **Question 9(b)**). One very common error was to give findings or conclusions when asked for advantages. This is illustrated by suggestions such as that 'It was an advantage because it showed us that ...'. Such answers are simply a statement of the fact that a useful conclusion could be drawn from the study; this is the purpose of research rather than an advantage of it.

In some cases the ability to apply knowledge was very good indeed (for example in **Question 12(a)**), but in other areas it was not so evident (for example in order to answer **Question 14(b)**).

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

Comments on specific questions

Section A

Question 1

- 1(a)** Many candidates gave simplistic, brief answers which could only earn 1 mark, such as 'They have an IV and a DV'. To earn full marks here, they needed to clarify what each was, e.g. that the IV is manipulated and the DV measured. Candidates also often referred to 'control', although rarely explained its role in the experimental method. A small number of candidates misunderstood what they were being asked to do and gave characteristics of the Mann et al. study, for example, 'they used videos'.
- (b)** Many answers to part **(b)** were not in sufficient detail, so could only earn 1 mark. Those candidates who were successful typically related their idea to lying. Some weaker candidates could show that they knew what a case study was but were not able to explain an advantage.

Question 2

- (a)** A common mistake here was to describe what happened in one or other of the interviews. A number of candidates incorrectly said that the participants were asked to talk to their families about the memory (rather than to *avoid* doing so). Nevertheless, there were also excellent answers, detailing either the need to think about the memory or to avoid talking to others, especially the family, about the study.

- (b) Relatively simple answers suggesting why it was important that they did not talk to their relatives were often able to earn full marks here. Other candidates gained full marks for effectively making the point that there would be potential to enhance the false memory.

Question 3

- (a) This question part was typically well answered, with candidates using all three examples of tests (IQ, eyes test and AQ) as well as the number of participants who were able to correctly identify gender to illustrate their answer.
- (b) This question part was not so well answered. Although many candidates were successful in answering correctly, using comparisons between the two groups, between genders or by referring to the negative correlation, marks were often lost as candidates either referred to “better”/“worse” (which is a value judgement, not a result), believed that group 3 scored higher, or gave the results of the eyes test rather than the AQ test.

Question 4

- (a) Most candidates knew this was a laboratory study but did not always correctly refer to the research method of a laboratory *experiment*. Marks were then sometimes earned for reference to controls. Some candidates mistakenly answered in terms of the experimental design.
- (b) There were many correct answers in this question part. However, common errors were to give answers which referred to findings (see above), or to give two brief advantages rather than one detailed one.

Question 5

- (a) Answers here were often weak. Most correct answers referred to stress, wanting to stop and wanting to know that the learner was OK. This was a very straightforward question although it revealed that many candidates were not familiar with the participants’ actual comments. Instead they guessed at what the participant (the ‘teacher’) may have said or used (incorrect) examples from watching films based on the study. Some candidates left this question blank.
- (b) This question part was answered well. Many candidates knew what self report data was, so were able to explain an advantage successfully.

Question 6

This question was generally well answered although responses were sometimes rather general, simply saying ‘to test helping’. There was also a tendency to list many variables that were tested or questions investigated, rather than to describe two reasons in detail. There was also some confusion over the belief that there were both male and female ‘victims’. Nevertheless, candidates seemed to have taken interest in the Kitty Genovese case, and they were able to report this well, and to discuss the reason why this was important in relation to the Piliavin et al. study.

Question 7

- (a) Most candidates were able to earn at least some credit here, for saying ‘maximum difference’, with many scoring full marks.
- (b) There were some excellent answers to part (b) of this question, with candidates demonstrating a sound understanding of the importance of the random allocation.

Question 8

Answers to this question demonstrated some good understanding of Freud’s ideas. The most common sources of evidence were the giraffe dream and the fantasy of his father becoming the grandfather. There was, however, occasional confusion in some candidates who simply described the Oedipus complex rather than the evidence, or who gave details of more than one source of evidence.

Question 9

- (a) This question was well answered by many candidates, although a common weak response was to simply refer to the preference for attractiveness. In other candidates, incorrect answers suggested guessing as they thought that the babies looked for longer at opposite sex faces.
- (b) Part (b) of this question was generally answered in relation to consent from parents although many other equally valid answers appeared in smaller numbers. There was a small minority of candidates who misunderstood the question and wrote about practical problems.

Question 10

- (a) This question part was not well answered by most candidates. There were, however, some candidates who knew Piaget's theory and could apply it to this question so answered well. It is important that candidates understand the essentials of the background for each core study. There were a significant number of candidates who muddled these studies with Bandura et al. and gave an answer relating to aggression.
- (b) Where candidates demonstrated an understanding of Piaget in part (a), they typically followed this up with a good response in relation to Nelson in part (b), but they were in the minority.

Question 11

- (a) There were few good answers for part (a) of this question. Many answers suggested that the candidate was unaware of the meaning of the term 'experimental design' and few recognised it as a repeated measures design. Even fewer could accurately describe this design. It is important that candidates have an understanding of basic research methodology. Of those candidates who correctly identified repeated measures, many went on to suggest that this meant repeatedly waking participants up to ask them questions. Consequently part (b) was not answered well.

Some candidates showed confusion with the concept of the 'research method' here, with candidates saying that this was an experiment. Some candidates also incorrectly gave conclusions here, rather than describing the design.

- (b) Part (b) of this question was not answered well, largely due to a lack of basic knowledge relating to experimental design. As this is a fundamental aspect of a key research method, it is vital that candidates can recognise and explain it. Some candidates incorrectly referred to conclusions here (as above).

Question 12

- (a) This question was well answered by many candidates. The most common response was volunteer sampling, with suggestions for newspaper or office advertisements. There was clear evidence of candidates utilising their knowledge from other core studies. Another common response was opportunity sampling, although there was some confusion with random sampling, as candidates often described opportunity sampling but called it random sampling.
- (b) In general, when part (a) was answered well, part (b) was also answered well. However, the confusion over random sampling led some candidates to make mistakes here. Nevertheless, when random sampling was correctly described in (a), there were good answers in relation to random sampling in (b).

Question 13

- (a) There were few correct responses here, and it appeared that many candidates did not understand or know the study well enough to answer well. Incorrect answers tended to include 'fresh air'.
- (b) Those candidates who answered correctly in (a) tended to score marks in (b) as well. However, incorrect answers, such as those about 'fresh air', tended to focus on controls rather than the importance of the smells being associated with bodies.

Question 14

- (a) Most candidates were able to give examples of data which could earn marks, although some candidates gave confused responses. Some candidates referred incorrectly to the numbers of apparent pseudo-patients 'identified' in the second study, which was quantitative data, but it needed to be collected by the pseudo-patients themselves to answer the question set.
- (b) Part (b) of this question produced a variety of good answers, with lack of demand characteristics being a common response. Some candidates gave conclusions from the study, a common weak response (see above).

Question 15

- (a) Responses to this question were typically weak, although there were some correct answers which demonstrated a clear comprehension of the question and recall of the study. A common incorrect response was to simply describe the aim of the study itself rather than the purpose of the specific question from the study.
- (b) Those candidates who gave correct answers in part (a) almost always scored full marks in part (b), demonstrating their understanding of both the method and the results of the study.

Section B

Question 16

The most popular choice here was the Haney, Banks and Zimbardo study although there were also good answers based on the Bandura et al. study. The small number of candidates using the Schachter and Singer study tended to produce very weak answers. The quality of responses was varied, with many candidates able to identify basic aspects of application or usefulness such as relevance to prison reforms for the Haney et al. study and parenting or television for the Bandura et al. study. Better essays were able to offer examples of these aspects from their chosen study and the strongest essays expanded on these basic ideas with a range of specific recommendations. Such essays were also more likely to be able to discuss reasons why their chosen study may *not* be very useful. Candidates who included no discussion of an absence of usefulness gained limited marks. At the lower end, these ideas were based on methodology but for better essays additional ideas were explored. Some candidates merely evaluated the study in general, earning marks only by coincidence, or just described the study, earning no marks. Such essays often illustrated that the candidates knew the studies well but did not have the skills to apply that knowledge to the question.

Question 17

This question produced a wider spread of choice across the three studies than **Question 17** and answers again demonstrated very good knowledge of the studies. The Thigpen and Cleckley study produced some good answers but also the weakest ones, for example with candidates often describing hypnosis as a psychometric test. Although answers were more balanced here, with candidates seemingly much clearer about the need to address strengths and weaknesses, there was again a tendency to consider these in relation to the study rather than using the study to illustrate an evaluation of psychometric tests in general.

PSYCHOLOGY

Paper 9698/21
Core Studies 2

Key messages

Section A

Question 1

It is important that candidates are made aware of the issues in psychology as some were unable to give a detailed definition of the laboratory experimental method in part **(a)**. Candidates should suggest a simple alternative to the original study in part **(b)** and give clear details of the procedure followed. In addition, candidates need to be aware of the requirements of this question as it does ask for an alternative to the original study. Extended evaluative points that make direct reference to the alternative idea are necessary in part **(c)** to achieve full marks.

Question 2

Centres should be aware that there are many different variations of the Milgram study and make sure they teach the variation that is in this syllabus. The Milgram study used is the 'remote victim condition' and not the 'victim response condition'. Many candidates incorrectly described the screams of the learner which were not from this variation.

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

Section B

Candidates must write extended responses in both part **(b)** and part **(c)**, many gave accurate responses that lacked depth or detailed responses that lacked focus on the question asked. Evidence must be given in part **(c)** to achieve higher marks.

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 number of candidates answered both questions in **Section B** and this has increased since the previous exam series. When a candidate did this they were awarded the mark for the best of the two questions (**Question 3** or **Question 4**). These candidates usually achieved very poorly.

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

Comments on specific questions

Section A

Question 1

(a) Most candidates achieved some marks for this question. Most were able to refer to control in some form and in appropriate context, but few described the use of an independent variable and measurement of the dependent variable. Many candidates used evaluative comments to answer this question such as a laboratory method being low in ecological validity or comparative comments with regards to other research methods such as the field, which did not answer the question and were not creditworthy. Very few candidates achieved full marks and very few mentioned causality.

(b) Most candidates were able to describe a procedure that used the field experiment method and was measuring infant facial preference. There were some very imaginative and detailed answers here. Many candidates were able to describe what was being measured and how. Venues for the field experiment tended to be a park, nursery, shopping mall, or the home of the infant.

Candidates often failed to include the sampling technique, which is important, alongside the number of participants. There was some confusion amongst candidates that the sample was the people who had their facial attractiveness assessed rather than the infants.

General attempts at designing an alternative to the Langlois et al. study discussed facial preference rather than infant preference for facial attractiveness and these responses gained limited credit, alongside any responses that used the laboratory experimental method or were unethical.

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

(c) The vast majority of candidates achieved marks in this question by providing some evaluative points. Most gave both methodological and ethical issues in their response. There were some excellent evaluation skills evident in terms of methodology with focus particularly on validity. For example the baby's gaze may have been affected by some other extraneous variable like a noise or a toy rather than the attractive or unattractive face. Ethical issues were described by most candidates and included mentioning issues of harm, consent and right to withdraw. These points were often quite brief.

Many candidates gained limited credit for evaluation that is simplistic but specific to the investigation. This was sometimes due to including one point or many points of evaluation in less detail with some in the context of their study from **Question 1(b)** whereas others were not linked.

Better answers demonstrated that the candidate could describe two or three evaluation issues, linked to their study and in some detail. The strongest answers demonstrated an ability to develop the issues raised in context.

Question 2

(a) The majority of candidates achieved credit in their answer to this question. Many candidates described the interaction with other people and how the environment/setting can influence human behaviour, but few referred to society or culture or group influences. A small minority of candidates used topic areas or studies as an example of the social approach.

(b) The vast majority of candidates achieved at least some credit for this question. The best answers described why the Milgram study is from the social approach, gave development of this point and then provided a clear link to how this was shown in the study (e.g. how obedience took place). Very few candidates managed to receive full marks but most were able to refer to the effect of the authority figure on the participant.

- (c) The vast majority of candidates achieved some credit in this section. Most were able to describe one strength and one weakness of the social approach and were able to give an example from the Milgram study. Many referred to control, reliability, ecological validity, usefulness and demand characteristics as appropriate points. Some responses focused on evaluation of the Milgram study rather than strengths and weaknesses of the social approach, using the Milgram study as an example.

Candidates need to describe two strengths and two weaknesses to achieve higher marks. Many did attempt to do this, but found it difficult to describe an appropriate second weakness (e.g. demand characteristics).

- (d) Many candidates answered appropriately for this question and were able to achieve high marks. Many candidates could identify at least two points about the extent to which the Milgram study can be considered ecologically valid.

There were some excellent descriptions of ways in which the Milgram study has both high and low ecological validity. Examples included how the participant received their own shock so this made the experiment real and how giving electric shocks to people is not a normal activity.

Some candidates failed to understand the term ecological validity and made reference to ethics and generalisability/representativeness, which was not creditworthy.

Candidates needed to give a consideration of ecological validity in detail and relevant to the Milgram study to be awarded high marks. This was usually achieved by candidates who outlined two or three points in detail rather than lots of smaller points in less detail.

Many candidates discussing the verbal responses made by the learner, which did not appear in the version of the study specified in the syllabus. For these candidates, this part of their answer did not receive any credit.

Section B

Question 3

- (a) The vast majority of candidates achieved some credit for this question and were aware that the individual differences approach investigates how individuals are unique or different. Very few candidates provided an example of a topic studied in this approach, such as mental health, or provided information related to differences in personality.
- (b) Many candidates attempted to answer the question and wrote about the behaviours investigated in each study, most scoring at least some credit for each study. Some candidates just described the procedure, but did get some marks for this description, where relevant.

Thigpen and Cleckley's study produced the best answers and some were able to give detailed descriptions of the behaviour of Eve investigated during the interviews with the therapists, as well as descriptions of the other behaviours investigated, such as her intelligence, personality, etc. For the Billington et al. study, many candidates described how empathising and systemising were investigated and a small minority of candidates specifically referred to the tests used in the study. For the Rosenhan study, some mentioned that the researcher investigated the reliability of diagnosis and the vast majority were able to briefly describe the pseudopatients' recording of the behaviour of staff and patients in the hospital. Many of the responses either lacked focus and/or were very brief.

- (c) Many candidates could identify one strength of investigating individual differences and many gave an example from a study to support their point. Some were able to identify two or three ideas but did not provide any evidence to support their points. The most popular strength discussed was the usefulness of the findings as well as the high level of ecological validity in most of the studies that investigate individual differences.

Many answers would benefit from the inclusion of more evidence to support points made. Some candidates referred to the same strength multiple times in their answer and gained limited credit.

Question 4

- (a) A few candidates gained full marks by providing a clear definition. Many candidates were unable to gain any credit for their response. A number mentioned either ecological validity or generalisability which was not creditworthy.
- (b) Most candidates achieved some credit for each study described in this question. They were able to describe an aspect of how the data were collected in each study, and/or give a correct detail of the study that was relevant. The studies by Haney, Banks and Zimbardo and Bandura et al. were described better than Tajfel. There was a tendency to provide irrelevant information from the study before addressing the question of how the data was collected. At times there was also too much focus on what data was collected rather than how it was collected. Other candidates described the study in too much detail which was not required and was not creditworthy.
- (c) Candidates need to identify and discuss three problems psychologists have when they try to make their research valid with clear reference to a core study for each point. Many were able to describe one or two problems such as ecological validity, ethics and generalisability/representativeness of the sample. Many did not link their responses to a study and achieved fewer marks.

PSYCHOLOGY

Paper 9698/22
Core Studies 2

Key messages

Section A

Question 1

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

Question 2

It is important for candidates to know how each study is linked to the methodology and data presentation, so for this examination, how the behaviourist perspective is linked to the Bandura et al. study. For part **(b)** candidates need to explain how the Bandura et al. study is an example of the behaviourist perspective. For part **(c)** candidates need to evaluate laboratory based studies using the Bandura et al. study as an example throughout, rather than just an evaluation of the Bandura et al. study. Also, to gain higher level marks candidates need to write about two strengths and two weaknesses as a minimum. For part **(d)** candidates need to be able to correctly appreciate the ethical issues raised by the Bandura et al. study.

Section B

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

General comments

Some candidates provided a range of excellent answers to many of the questions and could explain psychological terminology well, showing they had prepared themselves well for this paper. Only a few candidates gained top-end marks.

Time management appeared to be good for the majority of candidates. There was evidence that candidates who over-answered **Question 3(b) or 4(b)** wrote much shorter answers for **3(c) or 4(c)** as a result. This could also have been true for **Question 1(b)**. Candidates need to ensure they have enough time to answer all questions to the best of their ability.

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

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

Comments on specific questions

Section A

- (a) Many candidates could identify at least one type of observation. Types of data collection tended to not be mentioned by candidates. Some candidates provided evaluation of observations in their answer which is not what the question required, and could not be credited.
- (b) There was a wide variety of ideas given by candidates on how to examine smells and facial attractiveness as an observation. Many candidates could appropriately choose a sample and outline an appropriate sampling technique. Candidates usually did well outlining the *what* (which behaviours would be 'looked for') and the *where* (e.g. a university laboratory). Some candidates did not tackle the *how* clearly (the actual recording of the behaviours shown by the participants) so could not gain marks in the top band; simply writing 'behaviours were recorded' was seen as a major omission. There was a large number of answers where the candidate had not directly answered the question about 'smells and facial attractiveness' by designing studies just about smells. This type of question examines a candidate's ability to design a study that is practical and fully understands the potential methodology used by professional psychologists.
- (c) Many candidates could highlight one or two evaluative points about their own study designed in **Question 1(b)**. Common points made were about the sample used, and the unethical nature of a study about smells and facial attractiveness within a laboratory or field based study. A sizeable portion of candidates made a series of brief points linked to their own design to gain more marks. Some candidates evaluated aspects generically which gained limited credit as they had not linked them specifically to their own study.

Question 2

- (a) Many candidates clearly knew what the behaviourist perspective was and gained full marks. However, there was a sizeable amount of candidates who could not give an example or expand on a point already made to gain full marks.
- (b) Many candidates could explain how the Bandura et al. study was about learning/observing aggression to gain one mark. However, candidates found it difficult to then use examples from the Bandura et al. study to show how it is an example of behaviourism. Candidates tended to describe what Bandura et al. did without linking it to behaviourism.
- (c) Many candidates attempted this question and gained some credit. There was evidence of a reasonable minority of candidates evaluating the Bandura et al. study in general, answers needed to be linked to clear points about laboratory based studies. Many candidates could give some strengths and weaknesses of laboratory based studies but then failed to use the Bandura et al. study as examples of these strengths and weaknesses, so gained limited credit. There were many candidates attempting two strengths and two weaknesses where one of their evaluative points was about the Bandura et al. study *only* and not about laboratory based studies in general.
- (d) Many candidates could identify at least two ethical guidelines that were relevant to the Bandura et al. study. These included potential psychological harm and the problems related to debriefing young children. Candidates also attempted to argue about (informed) consent but this tended to be incorrect in terms of Bandura et al. obtaining it (it is not clear from the original paper). Some candidates however did argue the right to withdraw guideline well in terms of the experimenter being in the final room to ensure no child left.

Question 3

- (a) Many candidates could make reference to at least one aspect of a longitudinal study (e.g. over a time period) with an example to score maximum marks.
- (b) There were many very good answers to this question as candidates could pick out the necessary aspects of each study that showed how data were collected. Specific details about each study in terms of the actual data were crucial to gain the three marks per study. Rosenhan tended to be the study where candidates obtained the least marks. They tended to focus more on the actual results of the study rather than how data were collected. Many candidates could name at least two of the

measures taken by Thigpen and Cleckley, showing good knowledge of the study. Freud was covered well too, with many candidates being able to pick out the different ways in which he collected data about little Hans. However, there are candidates who give very long answers here that cover all of the study rather than having a focus on what the question is asking – in this case, data collection – and this is not an efficient use of time.

- (c) Some candidates could only manage brief answers here which could indicate they were not well prepared *or* that they had run out of time to write a more detailed response. Many candidates could at least outline some problems like longitudinal aspects of research and subject attrition and some then used a study to elaborate on the advantage. However, only a few candidates then went on to relate studies to all advantages. Some candidates are making the same points repeatedly and gaining only 3 marks in total – this is the maximum per problem written about.

Question 4

- (a) Many candidates could make reference to at least one and usually two aspects of qualitative data to score maximum marks.
- (b) Candidates appeared to know the three studies well but not in the context of qualitative *results*. The Haney, Banks and Zimbardo study was covered well in terms of what the participants went through and *how* data were collected but it was rare to see a candidate link this to actual results. The same applied to both the Milgram and Piliavin et al. studies. There are candidates who give very long answers here that cover all of the study rather than having a focus on what the question is asking – in this case, qualitative results – and this is not an efficient use of time.
- (c) A significant proportion of candidates could only manage brief answers here which could indicate they were not well prepared *or* that they had run out of time to write a more detailed response. Many candidates could outline one or two brief problems and this tended to be about subjectivity and experimenter bias. Only a minority of candidates could make three separate points and fewer could relate all to a study in order to gain the maximum of 3 marks per point made. As with **3(c)**, some candidates made the same point several times using different studies (especially about subjectivity/bias) but this could still only score the maximum of 3 marks for one well-made point with evidence.

PSYCHOLOGY

Paper 9698/23
Core Studies 2

Key messages

Section A

Question 1

It is important that candidates are made aware of the issues in psychology as some were unable to give a detailed definition of the laboratory experimental method in part **(a)**. Candidates should suggest a simple alternative to the original study in part **(b)** and give clear details of the procedure followed. In addition, candidates need to be aware of the requirements of this question as it does ask for an alternative to the original study. Extended evaluative points that make direct reference to the alternative idea are necessary in part **(c)** to achieve full marks.

Question 2

Centres should be aware that there are many different variations of the Milgram study and make sure they teach the variation that is in this syllabus. The Milgram study used is the 'remote victim condition' and not the 'victim response condition'. Many candidates incorrectly described the screams of the learner which were not from this variation.

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

Section B

Candidates must write extended responses in both part **(b)** and part **(c)**, many gave accurate responses that lacked depth or detailed responses that lacked focus on the question asked. Evidence must be given in part **(c)** to achieve higher marks.

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 number of candidates answered both questions in **Section B** and this has increased since the previous exam series. When a candidate did this they were awarded the mark for the best of the two questions (**Question 3** or **Question 4**). These candidates usually achieved very poorly.

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

Comments on specific questions

Section A

Question 1

(a) Most candidates achieved some marks for this question. Most were able to refer to control in some form and in appropriate context, but few described the use of an independent variable and measurement of the dependent variable. Many candidates used evaluative comments to answer this question such as a laboratory method being low in ecological validity or comparative comments with regards to other research methods such as the field, which did not answer the question and were not creditworthy. Very few candidates achieved full marks and very few mentioned causality.

(b) Most candidates were able to describe a procedure that used the field experiment method and was measuring infant facial preference. There were some very imaginative and detailed answers here. Many candidates were able to describe what was being measured and how. Venues for the field experiment tended to be a park, nursery, shopping mall, or the home of the infant.

Candidates often failed to include the sampling technique, which is important, alongside the number of participants. There was some confusion amongst candidates that the sample was the people who had their facial attractiveness assessed rather than the infants.

General attempts at designing an alternative to the Langlois et al. study discussed facial preference rather than infant preference for facial attractiveness and these responses gained limited credit, alongside any responses that used the laboratory experimental method or were unethical.

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

(c) The vast majority of candidates achieved marks in this question by providing some evaluative points. Most gave both methodological and ethical issues in their response. There were some excellent evaluation skills evident in terms of methodology with focus particularly on validity. For example the baby's gaze may have been affected by some other extraneous variable like a noise or a toy rather than the attractive or unattractive face. Ethical issues were described by most candidates and included mentioning issues of harm, consent and right to withdraw. These points were often quite brief.

Many candidates gained limited credit for evaluation that is simplistic but specific to the investigation. This was sometimes due to including one point or many points of evaluation in less detail with some in the context of their study from **Question 1(b)** whereas others were not linked.

Better answers demonstrated that the candidate could describe two or three evaluation issues, linked to their study and in some detail. The strongest answers demonstrated an ability to develop the issues raised in context.

Question 2

(a) The majority of candidates achieved credit in their answer to this question. Many candidates described the interaction with other people and how the environment/setting can influence human behaviour, but few referred to society or culture or group influences. A small minority of candidates used topic areas or studies as an example of the social approach.

(b) The vast majority of candidates achieved at least some credit for this question. The best answers described why the Milgram study is from the social approach, gave development of this point and then provided a clear link to how this was shown in the study (e.g. how obedience took place). Very few candidates managed to receive full marks but most were able to refer to the effect of the authority figure on the participant.

- (c) The vast majority of candidates achieved some credit in this section. Most were able to describe one strength and one weakness of the social approach and were able to give an example from the Milgram study. Many referred to control, reliability, ecological validity, usefulness and demand characteristics as appropriate points. Some responses focused on evaluation of the Milgram study rather than strengths and weaknesses of the social approach, using the Milgram study as an example.

Candidates need to describe two strengths and two weaknesses to achieve higher marks. Many did attempt to do this, but found it difficult to describe an appropriate second weakness (e.g. demand characteristics).

- (d) Many candidates answered appropriately for this question and were able to achieve high marks. Many candidates could identify at least two points about the extent to which the Milgram study can be considered ecologically valid.

There were some excellent descriptions of ways in which the Milgram study has both high and low ecological validity. Examples included how the participant received their own shock so this made the experiment real and how giving electric shocks to people is not a normal activity.

Some candidates failed to understand the term ecological validity and made reference to ethics and generalisability/representativeness, which was not creditworthy.

Candidates needed to give a consideration of ecological validity in detail and relevant to the Milgram study to be awarded high marks. This was usually achieved by candidates who outlined two or three points in detail rather than lots of smaller points in less detail.

Many candidates discussing the verbal responses made by the learner, which did not appear in the version of the study specified in the syllabus. For these candidates, this part of their answer did not receive any credit.

Section B

Question 3

- (a) The vast majority of candidates achieved some credit for this question and were aware that the individual differences approach investigates how individuals are unique or different. Very few candidates provided an example of a topic studied in this approach, such as mental health, or provided information related to differences in personality.
- (b) Many candidates attempted to answer the question and wrote about the behaviours investigated in each study, most scoring at least some credit for each study. Some candidates just described the procedure, but did get some marks for this description, where relevant.

Thigpen and Cleckley's study produced the best answers and some were able to give detailed descriptions of the behaviour of Eve investigated during the interviews with the therapists, as well as descriptions of the other behaviours investigated, such as her intelligence, personality, etc. For the Billington et al. study, many candidates described how empathising and systemising were investigated and a small minority of candidates specifically referred to the tests used in the study. For the Rosenhan study, some mentioned that the researcher investigated the reliability of diagnosis and the vast majority were able to briefly describe the pseudopatients' recording of the behaviour of staff and patients in the hospital. Many of the responses either lacked focus and/or were very brief.

- (c) Many candidates could identify one strength of investigating individual differences and many gave an example from a study to support their point. Some were able to identify two or three ideas but did not provide any evidence to support their points. The most popular strength discussed was the usefulness of the findings as well as the high level of ecological validity in most of the studies that investigate individual differences.

Many answers would benefit from the inclusion of more evidence to support points made. Some candidates referred to the same strength multiple times in their answer and gained limited credit.

Question 4

- (a) A few candidates gained full marks by providing a clear definition. Many candidates were unable to gain any credit for their response. A number mentioned either ecological validity or generalisability which was not creditworthy.
- (b) Most candidates achieved some credit for each study described in this question. They were able to describe an aspect of how the data were collected in each study, and/or give a correct detail of the study that was relevant. The studies by Haney, Banks and Zimbardo and Bandura et al. were described better than Tajfel. There was a tendency to provide irrelevant information from the study before addressing the question of how the data was collected. At times there was also too much focus on what data was collected rather than how it was collected. Other candidates described the study in too much detail which was not required and was not creditworthy.
- (c) Candidates need to identify and discuss three problems psychologists have when they try to make their research valid with clear reference to a core study for each point. Many were able to describe one or two problems such as ecological validity, ethics and generalisability/representativeness of the sample. Many did not link their responses to a study and achieved fewer marks.

PSYCHOLOGY

Paper 9698/31
Specialist Choices

Key messages

- Candidates should provide answers that equate to mark allocation, so an answer worth 2 marks should be short and an answer worth 8 marks should be correspondingly longer.
- Candidates should note that as this is a three hour examination and 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 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 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.
- It is essential that each question is correctly 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.

General comments

Responses must be legible. Many candidates used ink that leaked through onto the other side of the page and sometimes this made answers very difficult to read. Significant numbers of candidates used additional sheets of paper. It is helpful to Examiners if these are in the correct order and numbered.

A significant number of candidates appeared to be unable to answer some questions. This was particularly evident for the **Section B** ‘abnormal affect’ question for the Abnormality option, but also in some **Section A** questions (such as **Questions 5(a)** and **(b)**) and in a number of **Section C** questions (such as **Question 4(a)**). All syllabus sub-topics in candidates’ chosen options should be taught and revised, as any sub-topic can appear in an examination.

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)

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 did not apply to the question. For example, a candidate might write “intelligence is ecologically valid” or “pain is ethical” both of which have very little meaning. There are many issues that can be applied to every topic area and candidates are advised to think carefully about, and choose issues appropriate to, the topic area of the question.

Many answers would receive significantly higher marks if the difference between 'describe' and 'evaluate' is understood by candidates. **Section B** question part (a) will always be 'describe' and question part (b) will always be 'evaluate'. Evaluation is not simply additional description. Evaluation is a comment about what is good and what is not so good about the evidence that has been described in part (a). Evaluation requires a candidate to think and apply and not to just reproduce learning.

Those candidates who can evaluate can be divided into two types:

- those who evaluate using a number of evaluation issues in addition to the named issue (and these candidates score the highest marks);
- those who focus exclusively on the one named issue and gain limited marks, because one issue is not a range (as required by the mark scheme), or those candidates who exclude the named issue altogether (and also gain limited marks).

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).

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 including several superficial sentences about a number of different methods. Candidates need to show their methodological knowledge because many marks can be gained for application of this knowledge in this section.

Comments on specific questions

PSYCHOLOGY AND EDUCATION

- 1(a)** Most candidates scored full marks for this question part by giving a brief explanation of a learning difficulty or disability often supported with an example. A few candidates mentioned giftedness, but as this is neither a difficulty nor disability, it was not a creditable response.
- (b)** The most common disabilities referred to were attention deficit hyperactivity disorder, autistic spectrum disorder and dyslexia. A number of candidates wrote that the effects of dyslexia were dyscalculia and dyspraxia when they are not; they are types of disability in their own right.
- 2(a)** Many candidates wrote about Maslow and his hierarchy of needs but needed to link this to education, in order to answer the question set. Occasionally candidates also wrote about classical and operant conditioning (Pavlov and Skinner) without a mention of education, but this was less common. Answers at the top end of the mark range included these theories but then went on to say how they applied to education, often using examples from other topic areas, such as disruptive behaviour, to illustrate their answers. These candidates often made good use of attribution theory and learned helplessness which showed good understanding.
- (b)** The named evaluation issue here was 'different theories' and candidates had the choice of what other issues to include. Most candidates appeared to find evaluating 'different theories' straightforward. A few candidates adopted the strategy of evaluating one approach followed by another and ended up writing the same thing about each issue, commonly "it is reductionist", rather than comparing and contrasting different approaches. The quality of evaluation answers can be significantly improved for many candidates by planning the approach to the question.

- 3(a)** Most candidates scored some marks in relation to this question, but very few candidates scored marks at the top end of the range. Typical responses were to say “I would observe the child in the playground” without referring to any methodology about observations. The small number of responses simply said “I would use CCTV” without further explanation. Candidates should be aware that to score higher marks they need to show methodological knowledge. In this instance they could write about the type of observation, whether there will be one or two observers (to give inter-rater reliability), whether the observation will be time or event sampling, or what the response categories might be.
- (b)** Most candidates scored at least 1 mark for each type, even some common-sense responses such as ‘socio-economic status’ or ‘jealousy’ were creditable. Those candidates scoring full marks chose causes with a psychological basis, for example, jealousy in a person who has an ambivalent/anxious attachment style.
- 4(a)** This question gave candidates a free choice of method to consider how different types of intelligence were related. A logical approach would be to give a person two different tests and then correlate the resulting scores. Very few candidates suggested this, and those that did were not always able to expand on a simple statement, e.g. “I would correlate the scores”. Candidates can improve on their knowledge of correlation. Candidates are reminded that marks in this question part are awarded for methodological knowledge. Candidates were not always able to suggest two different types of intelligence; an IQ test and one of Gardner’s types could have been used, or any two of Gardner’s types could have been compared.
- (b)** Some candidates were able to describe different types of intelligence in this question part. A number of candidate listed different types, without description, and a few candidates included incorrect types. Some candidates wrote far more detail about the types than was needed for full marks.

PSYCHOLOGY AND HEALTH

- 5(a)** Many candidates answered the question well and often descriptions were supported with an appropriate example. Some candidates were unable to explain what a ‘life event’ was, and other candidates confused ‘life events’ with ‘daily hassles’.
- (b)** Many strong responses described the study by Holmes and Rahe in sufficient detail and included the underlying assumptions, examples from the questionnaire, how it was scored and the implications of a high score over a relatively short period of time. Some candidates were unable to describe a relevant study or incorrectly wrote about daily hassles.
- 6(a)** Many candidates wrote superb answers, with most of these scoring the maximum mark. Answers covered a wide range of appropriate aspects including methods (fear arousal and providing information) and studies conducted in schools, worksites and communities. A number of candidates included only one or two studies and these brief answers achieved no higher than mid-range marks. A few answers misinterpreted health promotion as adherence to medical requests and could not be credited.
- (b)** Some answers were impressive with the range and depth of argument, but other responses did not contain any evaluation. Some successful candidates had written about the studies by Tapper et al. and Walter in part **(a)** so had relevant examples to use to support the named issue for this question, ‘children’. Candidates were not always able to demonstrate knowledge of the syllabus sub-topic ‘health promotion in schools’ in part **(a)**, which would have provided examples to be used in this question part. Candidates are always advised to cover every aspect of the syllabus and are advised to cover three or more issues.

- 7(a)** Some candidates opted for a longitudinal approach, tracking adherence over twenty years or so, whilst others chose to do a cross sectional study, comparing levels of adherence in different age groups. In many answers, more demonstration of methodological knowledge was required, and terms were used without sufficient explanation of what they meant. If candidates choose to conduct an experiment (for example) then inclusion of IV and DV is essential. Writing “I would apply many controls” is appropriate, but it should be clearly explained what a control is, plus examples of what was controlled and why.
- (b)** The study named on the syllabus for rational non-adherence is that by Bulpitt and most candidates provided appropriate descriptions about what this study involved. This named study appears on the syllabus as an example, any other study of rational non-adherence would be acceptable, but no candidates wrote about an alternative study. A few candidates incorrectly wrote about customising treatment which could not be credited.
- 8(a)** In order to answer this question, candidates needed to understand the concept of token economy. Some candidates attempted to answer this question without a sufficient understanding, often scoring no marks. The strongest answers made appropriate comments not only about tokens being positive reinforcement, but also about the behaviourist approach on which the token system is based. A few candidates opted to conduct an experiment where, to determine the effectiveness of tokens on work safety, they gave an experimental group tokens and a control group nothing. This approach allowed these candidates to bring in appropriate methodological knowledge and to score high marks.
- (b)** Those candidates who understood what a token economy is, and wrote good answers in part **(a)**, were generally able to provide excellent answers to this question part when describing the study by Fox et al. Candidates described the problem, accidents in an open cast mine, explained how the token economy was introduced, and finally considered how work safety was significantly improved. Candidates who scored low marks in part **(a)** often could not describe any appropriate study in this question part. Candidates are advised to choose carefully which question to answer in this section.

PSYCHOLOGY AND ENVIRONMENT

- 9(a)** A significant number of candidates scored 1 mark or less because they failed to answer the question that was set, writing about ‘housing design’ rather than ‘community environmental design’, which are differentiated in the syllabus. This incorrect response also had implications for question part **(b)**. There were candidates who correctly explained that community environmental design concerns the design of public places to be used by any person in the community.
- (b)** The candidates who wrote correctly about community environmental design in part **(a)** most often wrote about the work of Whyte and Brower and frequently provided sufficient detail in answers for a maximum mark to be awarded. Credit was also awarded for the design of shopping malls because these are public places designed for use by anyone in the community. Candidates writing about housing design in part **(a)** often wrote about the work of Newman and his *housing* designs of Clason point, for example. These housing designs were not for public use and such answers scored no marks.

- 10(a)** Many candidates described what was meant by the terms density and crowding, often making the further distinction between social and spatial density. Stronger answers then considered a wide range of evidence, including studies involving both humans and animals, such as those by Christian, Calhoun, Lundberg and Evans and Wener. Many candidates went on to consider how crowding can be prevented, such as by modifying architecture, and how people can cope with its negative effects, by describing the studies by Langer and Saegert and Karlin et al. A few candidates showed a lack of understanding regarding some studies, for example assuming that the study by Evans and Wener was about coping with crowding when this study simply showed that 'three seat' designs, used on public transport, resulted in more stress than 'two seat' designs.
- (b)** The named issue here was 'ecological validity' and candidates were able to discuss this issue. For example the Calhoun animal study was conducted in a laboratory whereas the studies by Christian and all the other 'human' studies were not. Other relevant issues could be the reductionist nature of some studies; the use of physiological data (e.g. as used in the studies by Lundberg, and Evans and Wener). Top band answers included at least three different issues whereas those at the bottom end of the mark range included just one issue or provided 'general' evaluation. Notably many candidates assume that every study is reductionist and that it is something that is always negative. It is not, and candidates should also consider the positive aspects of conducting studies that are reductionist.
- 11(a)** Candidates had to use an interview method for this question. At the top end of the mark range candidates were able to distinguish between the type of interview (e.g. structured or unstructured) and also how the interview would be conducted (e.g. face-to-face or by telephone). These candidates often gave a few examples of questions and sometimes (although it was not required by the question) mentioned the ethics of conducting interviews with survivors of a traumatic event. A few candidates were not able to demonstrate any knowledge of the interview method.
- (b)** Candidates could often describe one study using non-interview methods, but were not always able to describe two. The most commonly described study, often done very well with good detail, was the laboratory experiment conducted by Mintz. For the second study a few candidates wrote about the social contagion *theory* proposed by Le Bon but this scored no marks because as a theory, no actual method was used. Some candidates wrote about the study on preparedness by Sattler which was appropriate. A few candidates also wrote about the theory of 'scripts' but this is again theory rather than an experimental or other study.
- 12(a)** Candidates had to investigate cultural differences in personal space and they were free to choose any method they wished. Some candidates decided to design a questionnaire, others opted to conduct a naturalistic observation. Most candidates decided to conduct a laboratory experiment using the stop-distance method. In many cases, the IV was people from different cultures and the DV was the distance at which they stopped from each other. Although these answers were often very good, answers could have been improved with the addition of more methodological terminology, such as application of an appropriate experimental design.
- (b)** There were many superb answers written in response to this question with many candidates scoring full marks. Descriptions of the study by Little were often thorough and included all the major components. There were many candidates who could not provide more than a few words on any study, or covered an incorrect study which could not be credited. Candidates are always advised to read questions thoroughly.

PSYCHOLOGY AND ABNORMALITY

- 13(a)** This question was generally answered well by most candidates. Weaker responses were simple statements such as “phobias are innate”, whereas answers scoring full marks provided elaboration in the form of detail or by referring to a relevant study such as that by Ost.
- (b)** Nearly all candidates could write an initial sentence about the biomedical/genetic explanation of phobias but could not always give expansion or an example to score further marks. Some candidates gave ambiguous examples and a common one was to say that “if the mother has a fear of snakes it is likely that the child will develop a fear of snakes”. This example could also illustrate the learning approach and is not a clear example that directly supports the requirements of the question. Most candidates referred to the study by Ost (1992) who found that those with a specific phobia for blood injuries had 60% of first degree relatives also having a specific phobia for blood injuries. There were some candidates who wrote about the work of Freud and little Hans which is not related to either a biomedical or genetic explanation.
- 14(a)** There were many candidates who knew the term, were well prepared, and wrote excellent answers which achieved full marks. A significant number of candidates did not understand the term ‘abnormal affect’, and rather than describing various forms of depression or mania, these candidates wrote about definitions of abnormality, such as deviation from social norms, for example. Others wrote about responses about models of abnormality.
- (b)** Some candidates extended their description from part **(a)**, rather than evaluating. A number of candidates only considered the named issue, in this case ‘competing explanations’, and received limited credit. Other candidates included the named issue as one of their three or four, but often mentioned the same issue for each study and so repeated what they wrote. Some candidates also mentioned issues in single sentences or just a few words, with no inclusion of a strength or weakness of that issue. For example, candidates would typically write “the biological approach is reductionist...” and a few sentences later they would write “Beck’s cognitive theory is reductionist”. Evaluation by evaluation issue is the most effective way to score high marks and this means starting with the issue, considering its advantages and disadvantages and then using approaches/studies as examples to illustrate those points. Many candidates consider reductionism as something negative when it is not.
- 15(a)** Candidates needed to demonstrate their knowledge of case studies and apply this to investigate Mrs Oliveira’s schizophrenia. Whilst many candidates did this, many others wrote nothing more than “I would interview her”, or “I will give her a questionnaire”, without including details such as the type of interview or type of questionnaire, and were unable to demonstrate methodological understanding. Some candidates suggested an interview but asked questions that were not related to schizophrenia.
- (b)** Candidates had to describe the underlying explanation for one treatment of schizophrenia. Some answers were superb, particularly those who chose medication, and description of the biochemical/medical approach and often in-depth knowledge of the chemistry of this approach was very good. Answers focusing on modifying the behavioural aspects of schizophrenia (such as the work by Paul and Lenz) also provided detailed answers.
- 16(a)** Some candidates made suggestions that were not based on psychological knowledge or showed no understanding of OCD. A few candidates suggested the use of medication, which wouldn’t target the specific behaviour identified in the question. Stronger responses had competent knowledge of cognitive-behaviour therapy and the strongest focused on how the cognitive aspects were directly related to the behavioural.
- (b)** This question required a description of the main features of the approach used to reduce the obsessive-compulsive disorder described in part **(a)**. Some candidates struggled to answer this question because their answer in part **(a)** had been vague, and those suggesting medication to reduce OCD in general wrote very little. Those suggesting the cognitive-behavioural approach in part **(a)** were often able to describe in detail the difference between cognition and the behavioural components and also provide a detailed description of the underlying theory.

Psychology and Organisations

- 17(a)** Some candidates were familiar with the term, and provided elaboration and/or example in support. Many candidates were unable to explain the term syllabus term 'performance appraisal techniques', appearing unfamiliar with it. No marks are awarded for reorganising the words of the question without demonstrating understanding of the term.
- (b)** Those candidates scoring no marks in part **(a)** appeared to guess, and wrote "I would interview them" and "I would give a questionnaire" without showing awareness of appropriate techniques or how appraisal is done in real life. A few candidates scored full marks by describing two appropriate techniques such as the use of check-lists, rating scales, the '360 technique' or 'BARS', the behaviourally anchored rating scale.
- 18(a)** Many candidates wrote excellent answers. Answers typically began with a distinction between intrinsic and extrinsic motivation, followed by a description of the work of Maslow, then others such as Alderfer, McClelland and Vroom. Some candidates also considered different reward systems. Weaker answers were often anecdotal and included nothing more than a list of common-sense factors or consisted of a very brief description of Maslow's hierarchy of needs.
- (b)** Although there were some superb answers in part **(a)** this rarely followed through into part **(b)**. Many candidates did not evaluate in the most effective way (see general comments and key messages). The named issue, of individual versus situational explanations, was included by most candidates, but not all. A few candidates only wrote about the named issue, for limited credit.
- 19(a)** This question required candidates to apply a questionnaire, and most candidates did this, with very few inappropriate methods being suggested. Candidates often moved beyond the basics when stating the type of questionnaire. Some included examples of questions, although these could have been more focused on flexitime. A few candidates mentioned a scoring or rating system and a small number of candidates mentioned how the data gathered could be analysed. A few candidates wrote about their sample, but overall, although describing the sample itself was good, very few wrote more than a few words about the sampling technique.
- (b)** Most candidates answering this question were not able to describe two ways in which the working week can be organised. Different types of shiftwork, compressed working weeks and flexitime all appear on the syllabus, so candidates need to have an understanding of these. If a candidate cannot provide any answer then the alternative question in **Section C** should be attempted instead. A few candidates could answer this question and most described the rapidly rotating patterns (e.g. the metropolitan and continental patterns) and a slowly rotating shift pattern and often scored full marks.
- 20(a)** Candidates had to investigate the popularity of a leader, and any method could be used. The most logical method to choose was either a questionnaire or interview with a rating scale to measure popularity. Most candidates opted for one of these, there were weak responses such as "I would use a questionnaire" without further detail. Candidates needed to include details such as "I would use a closed questionnaire with ten questions each using a 5 point scale ranging from 1 = not popular to 5 = very popular" to score most marks.
- (b)** Candidates had to describe a leader-member exchange model. Weaker responses were common-sense answers, which did not demonstrate psychological knowledge. Other candidates wrote about Tajfel's in-group and out-group but this does not really apply because one's own group is always the in-group whereas leader-member exchange is a different type of relationship. Answers at the top end of the mark range described either the suggested work or Scandura and Green in good detail, showing good understanding.

PSYCHOLOGY

Paper 9698/32
Specialist Choices

Key messages

- Candidates should provide answers that equate to mark allocation, so an answer worth 2 marks should be short and an answer worth 8 marks should be correspondingly longer.
- Candidates should note that as this is a three hour examination and 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 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 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.
- It is essential that each question is correctly 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.

General comments

Responses must be legible. Many candidates used ink that leaked through onto the other side of the page and sometimes this made answers very difficult to read. Significant numbers of candidates used additional sheets of paper. It is helpful to Examiners if these are in the correct order and numbered.

A significant number of candidates appeared to be unable to answer some questions. This was particularly evident for the **Section B** 'groups' question for the Organisations option, but also in some **Section A** questions (such as **Question 5(a)** and **(b)** on repeat prescriptions) and in a number of **Section C** questions. All syllabus sub-topics in candidates' chosen options should be taught and revised, as any sub-topic can appear in an examination.

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)

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 did not apply to the question. For example, a candidate might write "intelligence is ecologically valid" or "pain is ethical" both of which have very little meaning. There are many issues that can be applied to every topic area and candidates are advised to think carefully about, and choose issues appropriate to, the topic area of the question.

Many answers would receive significantly higher marks if the difference between 'describe' and 'evaluate' is understood by candidates. **Section B** question part (a) will always be 'describe' and question part (b) will always be 'evaluate'. Evaluation is not simply additional description. Evaluation is a comment about what is good and what is not so good about the evidence that has been described in part (a). Evaluation requires a candidate to think and apply and not to just reproduce learning.

Those candidates who can evaluate can be divided into two types:

- those who evaluate using a number of evaluation issues in addition to the named issue (and these candidates score the highest marks);
- those who focus exclusively on the one named issue and gain limited marks, because one issue is not a range (as required by the mark scheme), or those candidates who exclude the named issue altogether (and also gain limited marks).

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).

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 including several superficial sentences about a number of different methods. Candidates need to show their methodological knowledge because many marks can be gained for application of this knowledge in this section.

Comments on specific questions

PSYCHOLOGY AND EDUCATION

- 1(a)** Most candidates scored full marks by giving an appropriate explanation of preventive strategy, which is designed to stop a behaviour before it happens. A few answers muddled preventive with corrective which could not be credited.
- (b)** A very small number of candidates gave a preventive strategy, but most candidates provided a correct and appropriate answer. Some candidates focused on the stages of corrective techniques (e.g. Presland) whereas others suggested basic behaviourism and the application of positive reinforcement and positive or negative punishment. A small number of candidates confused these contingencies. Other candidates proposed the use of a token economy and often explained how it would work in a classroom.
- 2(a)** There were superb answers which included a brief history of the development of IQ tests, included a number of different theories of intelligence (such as Cattell, Gardner and Sternberg) and also included 'alternatives' to intelligence, such as creativity. There were many superb answers which showed excellent understanding and impressive detail. Some candidates were unable to demonstrate knowledge beyond the term IQ, and were not always able to explain how an IQ is calculated.
- (b)** The named evaluation issue here was 'predictive validity' and whilst many candidates knew the term, many others did not. Strong answers applied a wide range of different issues such as reliability, psychometric testing, reductionism and generalisations in addition to predictive validity. There were weaker answers which did not evaluate, which only included one issue or which did not include predictive validity.

- 3(a)** This question required candidates to use a longitudinal study, but the question did not specify which specific method to use over a period of time (because 'longitudinal' is not a method). Some candidates suggested using a questionnaire, others suggested a test conducted at regular intervals. All of these suggestions were appropriate. What determined marks was the methodological knowledge shown in answers. For example some candidates only identified aspects of methodology such as "I would use a random sample" or "I would match pairs" without expanding on how or why. A little additional detail explaining aspects such as these would result in stronger responses.
- (b)** Some candidates did not know Kolb's learning styles and others could only identify the four relevant words (for limited credit). Other candidates identified, described, elaborated and gave examples and resulting in full marks.
- 4(a)** Most candidates suggested conducting an experiment with one group of participants being given the CAT and another group being given nothing. Only a few candidates went on to identify that these were conditions of the independent variable, and, because a participant could not be in both the CAT and non-CAT group, that the design was independent groups. Most candidates needed to apply appropriate methodological terms to improve their answers. Some candidates struggled to identify the dependent variable, whereas others proposed some measure to assess the amount of change (and therefore the effectiveness of the CAT as the question requested).
- (b)** A significant number of candidates described the study conducted by Seligman on dogs from the Education option, which was creditable. Other candidates chose to describe the work of Dweck and on the whole this was very well done with some excellent descriptions being provided showing very good understanding.

PSYCHOLOGY AND HEALTH

- 5(a)** Many candidates were not able to demonstrate understanding of the syllabus term 'repeat prescription', and some gave a common-sense answer. Those candidates scoring full marks knew that a repeat prescription is a way in which adherence to medical requests can be measured and these candidates sometimes referred to the study by Sherman.
- (b)** Many candidates could not describe any study which used repeat prescriptions. Although the suggested example on the syllabus is the study by Sherman, candidates could have described an alternative study provided that it investigated repeat prescriptions. Description of studies looking at other forms of adherence such as that by Bulpitt on rational non-adherence, Chung and Naya on trackcap, or research on biochemical tests did not answer the question set and were not creditable. Many candidates described the Sherman study and wrote excellent answers which scored full marks.
- 6(a)** Many candidates wrote superb answers, with most of these scoring the maximum mark. Answers were well organised and covered a wide range of appropriate aspects listed on the syllabus. Most commonly the work by Argyle, McKinstry and Wang, McKinlay, Bryne and Long and Savage and Armstrong was well described. Some candidates went on to write about mis-using health services and described the work by Safer, Barlow and Durand, and Aleem and Ajarim. Weaker answers often included a narrower range of studies, described them in less detail and had inaccuracies in their description.
- (b)** Some answers provided an excellent range and depth of argument, and others did not include any evaluation. The named issue for this question was individual differences. Weaker answers included simple statements such as "there are individual differences" without elaboration, whereas strong responses dealt with individual differences as an issue, which involved using examples of studies and contrasting with what can be generalised (which minimises individual differences).

- 7(a)** Some candidates described a way in which pain could be measured without mentioning reliability or validity. Other candidates defined reliability and validity but did not mention pain. Some candidates attempted to answer the question simply by writing “and this makes it more reliable and valid” when what they were proposing would not make it so. The best answers used either test-retest or split half reliability and the most commonly considered way to assess validity was to use concurrent validity.
- (b)** The McGill (MPQ) Pain Questionnaire was commonly described, with candidates often showing good awareness of the different components of this test. Some candidates evaluated it, which was not required. Some candidates described the UAB pain observation checklist but often did not know enough about it to score full marks. The Paediatric Pain Questionnaire (PPQ) was also mentioned by some candidates but frequently answers were too brief to score full marks.
- 8(a)** The stem of the question gave a clear example of a study using a brain scanner (fMRI) to candidates, yet significant numbers did not appear to understand what a brain scan was or how it worked. For example, candidates suggested attaching electrodes to participants or conducting a field experiment with scans being done to measure the stress of driving a car. If a question in **Section C** is not understood then candidates should consider choosing the available alternative.
- (b)** The syllabus makes a clear distinction between measuring stress physiologically and measuring stress psychologically (using self-report questionnaires), but many candidates described two physiological measures. This often resulted in limited credit. Those candidates who described two appropriate measures often did so with ample detail. It is essential that candidates can distinguish between physiological measures (such as measures of blood pressure, heart rate, skin conductance, and blood and urine) and psychological measures (use of questionnaires).

PSYCHOLOGY AND ENVIRONMENT

- 9(a)** A number of candidates focused on *preventing* crowding (not allowing it to start) rather than *coping* (strategies to deal with it when it is happening), which could not be credited. Those candidates writing about coping often scored full marks, particularly when the answer included some elaboration and/or an example.
- (b)** Those answers concerned with *preventing* the effects of crowding could not be credited. This included any study on architecture (which is designed to prevent crowding). Some candidates also wrote about the study by Evans and Wener, but this was concerned with how seat design can cause crowding and not how people can cope. Many candidates addressed the question as intended and wrote some excellent answers, often describing the study by Langer and Saegert, or that by Karlin et al.
- 10(a)** Many candidates described what was meant by the terms ‘natural disaster’ and ‘technological catastrophe’, which was appropriate, and then provided multiple examples of each, for limited credit. Rather than a large number of examples, a range of descriptive points is required for higher marks, to show evidence that the psychology has been studied, and this should be done by quoting relevant theory and research. Answers at the top end of the mark range did this, quoting amongst others the work of Le Bon, Shank and Abelson, Hodgkinson and Stewart, Rubin et al., Mintz and Loftus.
- (b)** The named issue here was ‘the ethics of conducting studies’ and this was not done well by those candidates who provided examples of events only in part **(a)**. Those candidates who quoted appropriate studies in part **(a)**, such as that by Mintz, were able to raise many appropriate points in relation to ethics. Some of these candidates also provided superb answers because they considered ethics in relation to other issues; for example the more ethical a study is, the less true to real life it is likely to be. Many candidates wrote that because of the potential harm (such as injury and even death) a simulation study using deception could never be conducted.

- 11(a)** Candidates had to conduct a field experiment to investigate the effect of transport noise. A number of candidates used the word 'experiment' to mean 'study' and suggested studies that were not field experiments, which could not be credited. A few candidates even suggested conducting laboratory experiments which was also incorrect. Better answers did suggest field experiments but often failed to include details of an IV and DV, with a generalised procedure instead. Candidates need to answer the question as set, and should always apply as much methodological knowledge as they can.
- (b)** This question required a description of a study looking at the effect of noise on the performance of children, some weak responses covered studies that involved adults or that were not related to performance. The two most popular studies, which most candidates described very well, were those by Bronzaft and McCarthy (railway noise) and Haines et al. (airport noise).
- 12(a)** This question required candidates to think about how they could conduct a study using both humans and animals to investigate wayfinding using virtual reality. There were many very creative and interesting answers. Some answers were very good and showed some excellent understanding. The most logical suggestion was to create a virtual reality maze test that could be used on both humans (use of a virtual reality headset) and animals (such a rotating ball inside a 'cinema' environment).
- (b)** There were many superb answers written to this question with many candidates scoring full marks. There were excellent descriptions of a human study, most commonly those by Lynch or by Maguire, and there were excellent descriptions of an animal study with those by Jacobs and Linman, Capaldi and Walcott being most common. Some candidates wrote too much detail for three marks for each study and this time could have been applied more usefully elsewhere.

PSYCHOLOGY AND ABNORMALITY

- 13(a)** Many candidates knew the term and wrote excellent answers. Other candidates were not able to demonstrate understanding of 'type' or 'abnormal affect' and could not provide a creditable response. Incorrect answers often included definitions of abnormality and descriptions of schizophrenia and phobias. Abnormal affect is a major component of this syllabus option and it is essential that this term is understood.
- (b)** Candidates who knew the terms often distinguished between unipolar (depression) and bi-polar (mania and depression) and often wrote much more detail than was needed for two marks for each type. Other candidates were not able to demonstrate understanding of 'abnormal affect' and could not provide a creditable response.
- 14(a)** There were many superb answers written in response to this question on phobias by candidates who had been very well prepared. Candidates considered the different types of phobia (and gave many examples, such as agoraphobia), various explanations (cognitive, behavioural, genetic and psychoanalytic) and often provided a long list of appropriate treatments (including applied tension and systematic desensitisation).
- (b)** A number of candidates only considered the named issue of 'case studies' and received limited credit. Those candidates considering case studies in addition to other issues scored more marks, but these answers were often not well organised. For example, a candidate would evaluate the behavioural approach and refer to the little Albert study with a comment about it being a case study, then evaluate the psychoanalytic approach and make a reference to the little Hans study with a comment about it being a case study. A much more effective strategy is to include a paragraph on case studies, considering their strengths and weaknesses, and use the studies of little Albert and little Hans as supporting examples. This strategy should be repeated with another two or three issues.

- 15(a)** A small number of candidates considered the features of addiction in relation to drugs and alcohol exclusively, whereas most candidates considered the six features of addiction outlined by Griffiths which suggest that people can be addicted to anything. Some candidates listed the six features rather than describing them, for limited credit, whilst most candidates provided a detailed description and used several examples to illustrate the features.
- (b)** The best answers applied a method related to the features of addiction, outlined in part **(a)**. A range of different methods were suggested, such as observation, but also the use of questionnaires and interviews featured prominently. The best answers emphasised the methodology used; weaker answers focused less on methodology and more on internet applications such as Facebook.
- 16(a)** Most candidates were able to describe two definitions of abnormality. Many candidates described 'deviation from statistical norms' and were not able to supply supporting examples. Other definitions were 'deviation from social norms' which produced better examples, as did those candidates using 'deviation from ideal mental health', or those using 'failure to function adequately'.
- (b)** This question required an investigation using a questionnaire. Most candidates did this, but were not always able to include sufficient methodological knowledge. Candidates need to show knowledge of questionnaire type/design (e.g. open or closed), examples of questions (that clearly relate to defining abnormality), any rating scale that may be used, possibly where the questionnaire will be conducted, on whom, and how the answers will be scored. In addition to this, there are more general methodological issues that could be considered such as the sample, the sampling technique and how data will be analysed.

Psychology and Organisations

- 17(a)** Some candidates could demonstrate their understanding of the term 'sabotage' and some provided excellent definitions, very close to the original definition by Taylor and Walton. Other candidates were not able to demonstrate their understanding of the term and could not provide a creditable response. Candidates must be familiar with the syllabus terms for their chosen options.
- (b)** The vast majority of candidates described the appropriate two-factor theory by Herzberg. Most candidates could correctly outline 'motivators', although some answers needed to mention his 'hygiene' factors for higher marks.
- 18(a)** Weaker answers were often anecdotal, included a list of common-sense factors and needed to demonstrate psychological knowledge in order to be creditable. Stronger answers included well-described psychological studies across the range of relevant sub-topics including group development, team building, groupthink, group cohesiveness and group conflict.
- (b)** The named issue was individual versus situational explanations. There were candidates who did not include this issue at all or were unable to demonstrate clear understanding of it. Other candidates were very well prepared and discussed this issue in detail along with two or three others. Most candidates evaluated by study rather than by issue and whilst this scores some marks it is not the best exam technique (see general comments for further details).
- 19(a)** A number of candidates began this answer by *describing* the different types of interview. Whilst this information was often correct, it was not required by the question and could not be credited. Most candidates choose to use a questionnaire to gather data and many outlined the type of questionnaire they would conduct, often with a few sample questions in support. However, many candidates needed to include further detail, such as how closed questions would be scored or how the data from open ended questions would be analysed.
- (b)** The question required description of two ways in which people can apply for work, and no other aspect of the application/selection process. Many candidates answered the question correctly, writing about a curriculum vitae (CV) and application forms. However, many candidates answered incorrectly by focusing on different types of interviews and some wrote about selection decision-making processes.

- 20(a)** There was a free choice of method for this question, some candidates chose to conduct an interview with workers whilst others chose to distribute a questionnaire. Whilst these methods were creditable, many candidates also needed to cover the 'productivity' part of the question, which would not be easy to assess using a self-report. A few candidates suggested using an observation, but the majority decided to conduct an experiment with the DV being levels of happiness and levels of productivity.
- (b)** A few candidates incorrectly included monetary rewards in their answers, which could not be credited. Other candidates focused on non-monetary/intrinsic rewards such as praise, respect and recognition and often wrote very good answers. A few candidates wrote about job enrichment (and its specific types) and this approach also earned credit.

PSYCHOLOGY

Paper 9698/33
Specialist Choices

Key messages

- Candidates should provide answers that equate to mark allocation, so an answer worth 2 marks should be short and an answer worth 8 marks should be correspondingly longer.
- Candidates should note that as this is a three hour examination and 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 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 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.
- It is essential that each question is correctly 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.

General comments

Responses must be legible. Many candidates used ink that leaked through onto the other side of the page and sometimes this made answers very difficult to read. Significant numbers of candidates used additional sheets of paper. It is helpful to Examiners if these are in the correct order and numbered.

A significant number of candidates appeared to be unable to answer some questions. This was particularly evident for the **Section B** ‘abnormal affect’ question for the Abnormality option, but also in some **Section A** questions (such as **Questions 5(a)** and **(b)**) and in a number of **Section C** questions (such as **Question 4(a)**). All syllabus sub-topics in candidates’ chosen options should be taught and revised, as any sub-topic can appear in an examination.

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)

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 did not apply to the question. For example, a candidate might write “intelligence is ecologically valid” or “pain is ethical” both of which have very little meaning. There are many issues that can be applied to every topic area and candidates are advised to think carefully about, and choose issues appropriate to, the topic area of the question.

Many answers would receive significantly higher marks if the difference between ‘describe’ and ‘evaluate’ is understood by candidates. **Section B** question part (a) will always be ‘describe’ and question part (b) will always be ‘evaluate’. Evaluation is not simply additional description. Evaluation is a comment about what is good and what is not so good about the evidence that has been described in part (a). Evaluation requires a candidate to think and apply and not to just reproduce learning.

Those candidates who can evaluate can be divided into two types:

- those who evaluate using a number of evaluation issues in addition to the named issue (and these candidates score the highest marks);
- those who focus exclusively on the one named issue and gain limited marks, because one issue is not a range (as required by the mark scheme), or those candidates who exclude the named issue altogether (and also gain limited marks).

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).

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 including several superficial sentences about a number of different methods. Candidates need to show their methodological knowledge because many marks can be gained for application of this knowledge in this section.

Comments on specific questions

PSYCHOLOGY AND EDUCATION

- 1(a)** Most candidates scored full marks for this question part by giving a brief explanation of a learning difficulty or disability often supported with an example. A few candidates mentioned giftedness, but as this is neither a difficulty nor disability, it was not a creditable response.
- (b)** The most common disabilities referred to were attention deficit hyperactivity disorder, autistic spectrum disorder and dyslexia. A number of candidates wrote that the effects of dyslexia were dyscalculia and dyspraxia when they are not; they are types of disability in their own right.
- 2(a)** Many candidates wrote about Maslow and his hierarchy of needs but needed to link this to education, in order to answer the question set. Occasionally candidates also wrote about classical and operant conditioning (Pavlov and Skinner) without a mention of education, but this was less common. Answers at the top end of the mark range included these theories but then went on to say how they applied to education, often using examples from other topic areas, such as disruptive behaviour, to illustrate their answers. These candidates often made good use of attribution theory and learned helplessness which showed good understanding.
- (b)** The named evaluation issue here was ‘different theories’ and candidates had the choice of what other issues to include. Most candidates appeared to find evaluating ‘different theories’ straightforward. A few candidates adopted the strategy of evaluating one approach followed by another and ended up writing the same thing about each issue, commonly “it is reductionist”, rather than comparing and contrasting different approaches. The quality of evaluation answers can be significantly improved for many candidates by planning the approach to the question.

- 3(a)** Most candidates scored some marks in relation to this question, but very few candidates scored marks at the top end of the range. Typical responses were to say “I would observe the child in the playground” without referring to any methodology about observations. The small number of responses simply said “I would use CCTV” without further explanation. Candidates should be aware that to score higher marks they need to show methodological knowledge. In this instance they could write about the type of observation, whether there will be one or two observers (to give inter-rater reliability), whether the observation will be time or event sampling, or what the response categories might be.
- (b)** Most candidates scored at least 1 mark for each type, even some common-sense responses such as ‘socio-economic status’ or ‘jealousy’ were creditable. Those candidates scoring full marks chose causes with a psychological basis, for example, jealousy in a person who has an ambivalent/anxious attachment style.
- 4(a)** This question gave candidates a free choice of method to consider how different types of intelligence were related. A logical approach would be to give a person two different tests and then correlate the resulting scores. Very few candidates suggested this, and those that did were not always able to expand on a simple statement, e.g. “I would correlate the scores”. Candidates can improve on their knowledge of correlation. Candidates are reminded that marks in this question part are awarded for methodological knowledge. Candidates were not always able to suggest two different types of intelligence; an IQ test and one of Gardner’s types could have been used, or any two of Gardner’s types could have been compared.
- (b)** Some candidates were able to describe different types of intelligence in this question part. A number of candidate listed different types, without description, and a few candidates included incorrect types. Some candidates wrote far more detail about the types than was needed for full marks.

PSYCHOLOGY AND HEALTH

- 5(a)** Many candidates answered the question well and often descriptions were supported with an appropriate example. Some candidates were unable to explain what a ‘life event’ was, and other candidates confused ‘life events’ with ‘daily hassles’.
- (b)** Many strong responses described the study by Holmes and Rahe in sufficient detail and included the underlying assumptions, examples from the questionnaire, how it was scored and the implications of a high score over a relatively short period of time. Some candidates were unable to describe a relevant study or incorrectly wrote about daily hassles.
- 6(a)** Many candidates wrote superb answers, with most of these scoring the maximum mark. Answers covered a wide range of appropriate aspects including methods (fear arousal and providing information) and studies conducted in schools, worksites and communities. A number of candidates included only one or two studies and these brief answers achieved no higher than mid-range marks. A few answers misinterpreted health promotion as adherence to medical requests and could not be credited.
- (b)** Some answers were impressive with the range and depth of argument, but other responses did not contain any evaluation. Some successful candidates had written about the studies by Tapper et al. and Walter in part **(a)** so had relevant examples to use to support the named issue for this question, ‘children’. Candidates were not always able to demonstrate knowledge of the syllabus sub-topic ‘health promotion in schools’ in part **(a)**, which would have provided examples to be used in this question part. Candidates are always advised to cover every aspect of the syllabus and are advised to cover three or more issues.

- 7(a)** Some candidates opted for a longitudinal approach, tracking adherence over twenty years or so, whilst others chose to do a cross sectional study, comparing levels of adherence in different age groups. In many answers, more demonstration of methodological knowledge was required, and terms were used without sufficient explanation of what they meant. If candidates choose to conduct an experiment (for example) then inclusion of IV and DV is essential. Writing “I would apply many controls” is appropriate, but it should be clearly explained what a control is, plus examples of what was controlled and why.
- (b)** The study named on the syllabus for rational non-adherence is that by Bulpitt and most candidates provided appropriate descriptions about what this study involved. This named study appears on the syllabus as an example, any other study of rational non-adherence would be acceptable, but no candidates wrote about an alternative study. A few candidates incorrectly wrote about customising treatment which could not be credited.
- 8(a)** In order to answer this question, candidates needed to understand the concept of token economy. Some candidates attempted to answer this question without a sufficient understanding, often scoring no marks. The strongest answers made appropriate comments not only about tokens being positive reinforcement, but also about the behaviourist approach on which the token system is based. A few candidates opted to conduct an experiment where, to determine the effectiveness of tokens on work safety, they gave an experimental group tokens and a control group nothing. This approach allowed these candidates to bring in appropriate methodological knowledge and to score high marks.
- (b)** Those candidates who understood what a token economy is, and wrote good answers in part **(a)**, were generally able to provide excellent answers to this question part when describing the study by Fox et al. Candidates described the problem, accidents in an open cast mine, explained how the token economy was introduced, and finally considered how work safety was significantly improved. Candidates who scored low marks in part **(a)** often could not describe any appropriate study in this question part. Candidates are advised to choose carefully which question to answer in this section.

PSYCHOLOGY AND ENVIRONMENT

- 9(a)** A significant number of candidates scored 1 mark or less because they failed to answer the question that was set, writing about ‘housing design’ rather than ‘community environmental design’, which are differentiated in the syllabus. This incorrect response also had implications for question part **(b)**. There were candidates who correctly explained that community environmental design concerns the design of public places to be used by any person in the community.
- (b)** The candidates who wrote correctly about community environmental design in part **(a)** most often wrote about the work of Whyte and Brower and frequently provided sufficient detail in answers for a maximum mark to be awarded. Credit was also awarded for the design of shopping malls because these are public places designed for use by anyone in the community. Candidates writing about housing design in part **(a)** often wrote about the work of Newman and his *housing* designs of Clason point, for example. These housing designs were not for public use and such answers scored no marks.

- 10(a)** Many candidates described what was meant by the terms density and crowding, often making the further distinction between social and spatial density. Stronger answers then considered a wide range of evidence, including studies involving both humans and animals, such as those by Christian, Calhoun, Lundberg and Evans and Wener. Many candidates went on to consider how crowding can be prevented, such as by modifying architecture, and how people can cope with its negative effects, by describing the studies by Langer and Saegert and Karlin et al. A few candidates showed a lack of understanding regarding some studies, for example assuming that the study by Evans and Wener was about coping with crowding when this study simply showed that 'three seat' designs, used on public transport, resulted in more stress than 'two seat' designs.
- (b)** The named issue here was 'ecological validity' and candidates were able to discuss this issue. For example the Calhoun animal study was conducted in a laboratory whereas the studies by Christian and all the other 'human' studies were not. Other relevant issues could be the reductionist nature of some studies; the use of physiological data (e.g. as used in the studies by Lundberg, and Evans and Wener). Top band answers included at least three different issues whereas those at the bottom end of the mark range included just one issue or provided 'general' evaluation. Notably many candidates assume that every study is reductionist and that it is something that is always negative. It is not, and candidates should also consider the positive aspects of conducting studies that are reductionist.
- 11(a)** Candidates had to use an interview method for this question. At the top end of the mark range candidates were able to distinguish between the type of interview (e.g. structured or unstructured) and also how the interview would be conducted (e.g. face-to-face or by telephone). These candidates often gave a few examples of questions and sometimes (although it was not required by the question) mentioned the ethics of conducting interviews with survivors of a traumatic event. A few candidates were not able to demonstrate any knowledge of the interview method.
- (b)** Candidates could often describe one study using non-interview methods, but were not always able to describe two. The most commonly described study, often done very well with good detail, was the laboratory experiment conducted by Mintz. For the second study a few candidates wrote about the social contagion *theory* proposed by Le Bon but this scored no marks because as a theory, no actual method was used. Some candidates wrote about the study on preparedness by Sattler which was appropriate. A few candidates also wrote about the theory of 'scripts' but this is again theory rather than an experimental or other study.
- 12(a)** Candidates had to investigate cultural differences in personal space and they were free to choose any method they wished. Some candidates decided to design a questionnaire, others opted to conduct a naturalistic observation. Most candidates decided to conduct a laboratory experiment using the stop-distance method. In many cases, the IV was people from different cultures and the DV was the distance at which they stopped from each other. Although these answers were often very good, answers could have been improved with the addition of more methodological terminology, such as application of an appropriate experimental design.
- (b)** There were many superb answers written in response to this question with many candidates scoring full marks. Descriptions of the study by Little were often thorough and included all the major components. There were many candidates who could not provide more than a few words on any study, or covered an incorrect study which could not be credited. Candidates are always advised to read questions thoroughly.

PSYCHOLOGY AND ABNORMALITY

- 13(a)** This question was generally answered well by most candidates. Weaker responses were simple statements such as “phobias are innate”, whereas answers scoring full marks provided elaboration in the form of detail or by referring to a relevant study such as that by Ost.
- (b)** Nearly all candidates could write an initial sentence about the biomedical/genetic explanation of phobias but could not always give expansion or an example to score further marks. Some candidates gave ambiguous examples and a common one was to say that “if the mother has a fear of snakes it is likely that the child will develop a fear of snakes”. This example could also illustrate the learning approach and is not a clear example that directly supports the requirements of the question. Most candidates referred to the study by Ost (1992) who found that those with a specific phobia for blood injuries had 60% of first degree relatives also having a specific phobia for blood injuries. There were some candidates who wrote about the work of Freud and little Hans which is not related to either a biomedical or genetic explanation.
- 14(a)** There were many candidates who knew the term, were well prepared, and wrote excellent answers which achieved full marks. A significant number of candidates did not understand the term ‘abnormal affect’, and rather than describing various forms of depression or mania, these candidates wrote about definitions of abnormality, such as deviation from social norms, for example. Others wrote about responses about models of abnormality.
- (b)** Some candidates extended their description from part **(a)**, rather than evaluating. A number of candidates only considered the named issue, in this case ‘competing explanations’, and received limited credit. Other candidates included the named issue as one of their three or four, but often mentioned the same issue for each study and so repeated what they wrote. Some candidates also mentioned issues in single sentences or just a few words, with no inclusion of a strength or weakness of that issue. For example, candidates would typically write “the biological approach is reductionist...” and a few sentences later they would write “Beck’s cognitive theory is reductionist”. Evaluation by evaluation issue is the most effective way to score high marks and this means starting with the issue, considering its advantages and disadvantages and then using approaches/studies as examples to illustrate those points. Many candidates consider reductionism as something negative when it is not.
- 15(a)** Candidates needed to demonstrate their knowledge of case studies and apply this to investigate Mrs Oliveira’s schizophrenia. Whilst many candidates did this, many others wrote nothing more than “I would interview her”, or “I will give her a questionnaire”, without including details such as the type of interview or type of questionnaire, and were unable to demonstrate methodological understanding. Some candidates suggested an interview but asked questions that were not related to schizophrenia.
- (b)** Candidates had to describe the underlying explanation for one treatment of schizophrenia. Some answers were superb, particularly those who chose medication, and description of the biochemical/medical approach and often in-depth knowledge of the chemistry of this approach was very good. Answers focusing on modifying the behavioural aspects of schizophrenia (such as the work by Paul and Lenz) also provided detailed answers.
- 16(a)** Some candidates made suggestions that were not based on psychological knowledge or showed no understanding of OCD. A few candidates suggested the use of medication, which wouldn’t target the specific behaviour identified in the question. Stronger responses had competent knowledge of cognitive-behaviour therapy and the strongest focused on how the cognitive aspects were directly related to the behavioural.
- (b)** This question required a description of the main features of the approach used to reduce the obsessive-compulsive disorder described in part **(a)**. Some candidates struggled to answer this question because their answer in part **(a)** had been vague, and those suggesting medication to reduce OCD in general wrote very little. Those suggesting the cognitive-behavioural approach in part **(a)** were often able to describe in detail the difference between cognition and the behavioural components and also provide a detailed description of the underlying theory.

Psychology and Organisations

- 17(a)** Some candidates were familiar with the term, and provided elaboration and/or example in support. Many candidates were unable to explain the term syllabus term 'performance appraisal techniques', appearing unfamiliar with it. No marks are awarded for reorganising the words of the question without demonstrating understanding of the term.
- (b)** Those candidates scoring no marks in part **(a)** appeared to guess, and wrote "I would interview them" and "I would give a questionnaire" without showing awareness of appropriate techniques or how appraisal is done in real life. A few candidates scored full marks by describing two appropriate techniques such as the use of check-lists, rating scales, the '360 technique' or 'BARS', the behaviourally anchored rating scale.
- 18(a)** Many candidates wrote excellent answers. Answers typically began with a distinction between intrinsic and extrinsic motivation, followed by a description of the work of Maslow, then others such as Alderfer, McClelland and Vroom. Some candidates also considered different reward systems. Weaker answers were often anecdotal and included nothing more than a list of common-sense factors or consisted of a very brief description of Maslow's hierarchy of needs.
- (b)** Although there were some superb answers in part **(a)** this rarely followed through into part **(b)**. Many candidates did not evaluate in the most effective way (see general comments and key messages). The named issue, of individual versus situational explanations, was included by most candidates, but not all. A few candidates only wrote about the named issue, for limited credit.
- 19(a)** This question required candidates to apply a questionnaire, and most candidates did this, with very few inappropriate methods being suggested. Candidates often moved beyond the basics when stating the type of questionnaire. Some included examples of questions, although these could have been more focused on flexitime. A few candidates mentioned a scoring or rating system and a small number of candidates mentioned how the data gathered could be analysed. A few candidates wrote about their sample, but overall, although describing the sample itself was good, very few wrote more than a few words about the sampling technique.
- (b)** Most candidates answering this question were not able to describe two ways in which the working week can be organised. Different types of shiftwork, compressed working weeks and flexitime all appear on the syllabus, so candidates need to have an understanding of these. If a candidate cannot provide any answer then the alternative question in **Section C** should be attempted instead. A few candidates could answer this question and most described the rapidly rotating patterns (e.g. the metropolitan and continental patterns) and a slowly rotating shift pattern and often scored full marks.
- 20(a)** Candidates had to investigate the popularity of a leader, and any method could be used. The most logical method to choose was either a questionnaire or interview with a rating scale to measure popularity. Most candidates opted for one of these, there were weak responses such as "I would use a questionnaire" without further detail. Candidates needed to include details such as "I would use a closed questionnaire with ten questions each using a 5 point scale ranging from 1 = not popular to 5 = very popular" to score most marks.
- (b)** Candidates had to describe a leader-member exchange model. Weaker responses were common-sense answers, which did not demonstrate psychological knowledge. Other candidates wrote about Tajfel's in-group and out-group but this does not really apply because one's own group is always the in-group whereas leader-member exchange is a different type of relationship. Answers at the top end of the mark range described either the suggested work or Scandura and Green in good detail, showing good understanding.