PSYCHOLOGY

Paper 9990/12

Paper 1 Approaches, Issues and Debates

Key messages

Candidates need to know all components of the study as listed in the syllabus. Questions can be asked about any part of a study.

Candidates need to read the whole question carefully to ensure that their responses are fulfilling the demands of each one. For example, the question may require data or a named issue to be included. To achieve full marks, these need to be correctly present in their responses. The essay (final question) requires four evaluation points to be in depth (two strengths and two weaknesses) with at least one of these about the named issue. Credit is limited if the named issue is omitted or just described.

Candidates need to be careful about how they are presenting the results of studies. For example, they need to know if the results are about how many participants performed a task correctly or on how many trials the participant was correct. This can have a large impact on the interpretation of results and whether a response can gain credit.

Candidates also need to engage with any stimulus material presented in a question (for example, a novel situation) to ensure they can access all available marks. In addition, when a question refers to 'in this study' the answer requires contextualisation.

Candidates need to be able to consider real-world applications for all core studies. To show understanding, answers need to tell the examiner what the application is, based on the particular core study, and then how this could be achieved.

Candidates need to appreciate the difference between a result and a conclusion. The former is factual and based on collected data. The latter is a generic comment based on the results reported in any core study.

Candidates also need to know the set procedure of studies in the order presented in the original journal article. Questions can be based around just *part* of a procedure and the candidate must be able to produce an answer that is directed and concise rather than writing about the whole of the procedure.

Questions about the 'psychology being investigated' require answers that are generically describing concepts and theories that are central to the study. Minimal credit can be awarded for answers which describe the processes of the study itself rather than a generic description of a theory or a concept.

There is enough time for answers to be planned to ensure that the response given by a candidate is focused on the demands of each question.

General comments

The marks achieved by the candidates sitting this examination covered a wide spread of possible marks. Some candidates provided a range of excellent answers to many of the questions and could explain psychological terminology well providing evidence that they were prepared for the examination.

Stronger overall responses followed the demands of each question with explicit use of psychological terminology and logical, well planned answers in evidence. Appropriate examples were used from studies when the question expected it and there was evidence of candidates being able to apply their knowledge to real-world behaviours in terms of what and how.

Comments on specific questions

Question 1

(a) Many responses to this question were correct (REM). Some responses named a different stage of sleep which could not gain credit.

(b) There were many correct responses to this question. Stronger responses could clearly outline a dream reported that had vertical eye movements. The most common response was climbing up a ladder, looking up and down. Some responses described a horizontal eye movement dream so could not gain credit.

(c) There were some correct responses to this question. Popular correct responses focused on judging duration of REM and eye movements corresponding to content of dream. However, some responses gave a result from the study and therefore could not gain any credit. A result is factual, based on presented data whereas a conclusion is the interpretation of the results.

Question 2

(a) There were some correct responses with popular choices being race, sex and location of the passengers. Some responses gave the results reported from the critical area or focused on the adjacent area, which could not be credited.

(b) (i) There were many correct responses with some detailed descriptions of diffusion of responsibility that gained maximum credit. The process of it being active and linked to group size were popular explanations. However, there were some responses that used a reverse argument of fewer people = less help which is not diffusion of responsibility, or described the process as being passive when it is active.

(ii) The responses to this question were mixed. Some responses could clearly describe the pattern of result and use one correct piece of data to gain maximum credit. The most popular correct response was about larger groups responding faster than smaller groups. Some responses confused the two groups and gave the reverse result. Other responses described incorrect results based around larger groups being more likely to help – this is not correct – the larger groups simply responded faster. It is important for candidates to know the key results from all core studies and how data were presented.

Question 3

For responses to gain maximum credit, real-world applications had to explain what the application was, how it would be achieved based on the study by Bandura et al. and be ethical. Popular creditworthy choices included helping to reduce aggression in problem children and helping TV programmers ensure appropriate age-related content for children to view. The 'how' part of the question was weaker in most responses. It should be made clear by the candidate how the application can be achieved or used to explain an event. The candidate should not assume that the examiner will make the link. It must be explicitly made by the candidate. There were very few responses that were unethical.

Question 4

(a) Some responses could clearly identify two characteristics of the matched scenes. Some responses confused the characteristics with controls used, such as the amount of time the scenes were shown for. These answers could not gain credit.

(b) Most responses could clearly describe the sample of participants used. Popular creditworthy points included the sample size, that they were all female, volunteer and were right-handed.

(c) Many responses could outline an assumption of the biological approach that could be used to explain why the study by Canli et al. is from that approach. Stronger responses could explicitly then explain, using an example from the study, why the assumption links to that aspect of the study by Canli et al. For example, the assumption of how the brain affects cognition/behaviour/feelings is correct, so a full-mark response would then use the example of how it was the amygdala affecting the memory of emotional stimuli/pictures.

Question 5

The majority of responses to this question could only describe one or two aspects of the specified section of procedure from the study by Milgram. Many responses described the fact that the responses were predetermined and how there was no communication from the learner until 300 v. However, many responses could then go no further in describing this part of the procedure or gave answers that were from the procedure of a *different* Milgram study. Candidates need to know the full procedure for all core studies listed in the syllabus to avoid confusion with variations or follow-up studies.

Question 6

To improve responses to this type of question, candidates must focus on the general psychology being investigated in the study by Pepperberg. One mark is available for an example from the study, but the remainder must come from general psychology. Stronger responses could clearly outline aspects like operant conditioning or the use of social learning. However, many responses described the study without highlighting what general psychology was being investigated. Candidates need to be prepared for this aspect of the syllabus by being able to describe generic psychology and not the specific aims of the study as this can be examined through questions specifically about the aim and procedure of a core study.

Question 7

(a) Most responses could clearly outline one aim of the study by Yamamoto et al. Popular choices included investigating targeted helping and/or altruism. Some responses provided a finding from the study rather than an aim so could not gain credit.

(b) There was a range of well-chosen methodological weaknesses for the study by Yamamoto et al. Popular choices included the issues with tasks lacking mundane realism and sample size. For the latter, there were strong responses explaining how the mother-juvenile pairings may be qualitatively different to other chimpanzees either in a laboratory or in the wild due to them already being used extensively in studies. There were some responses that focused on ethical weaknesses so could not gain credit. To improve, candidates must ensure that once an appropriate methodological weakness has been chosen that it is then explicitly linked to the study by Yamamoto et al. by example, to address the 'of this study' in the question.

(c) There were many examples of real-world applications based on the findings of Yamamoto et al. Popular choices included using a similar procedure to aid the teaching of helping behaviour in the classroom. Many responses could outline what the application would be, but a significant amount could not explain how it would be implemented to gain both available marks. Responses that simply state 'using the same procedures as Yamamoto' cannot be awarded the 'how' mark. This part must be explicitly outlined by the candidate.

Question 8

(a) Most responses could outline the nature versus nurture debate. Some responses could provide a suitable example for either side of the debate, either from everyday behaviours or from a core study. Popular choices were examples from Bandura et al., Pepperberg and Dement and Kleitman. There were several responses that were tautological which could not gain credit. By this, the candidate would state that the nature explanation was about the nature of the individual and that the nurture explanation was about the nurturing of a person. Issues and debates need to be outlined without using the term to outline the term to gain maximum credit. In addition, the candidate needs to explicitly label which is the definition of nature and which is the definition of nurture, to clearly demonstrate their understanding. An examiner cannot provide the label for the candidate.

(b) Stronger responses could clearly choose relevant results and use them to explain why the result supported the nature and nurture sides of the debate. A popular choice for the nature side was the pulse rates increasing for all participants in an epinephrine condition. This could then be explained using ideas like us being hard wired to react to that chemical in this way. A popular choice for the nurture side was the EPI-MIS group being happier as they followed the situation that they were in – they have learned to be happy! However, there were responses that could not provide an actual result from the study and/or use it to explain why the result was supporting either side of the debate.

Question 9

The strongest responses evaluated the study by Laney et al. in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of independent measures. Common choices included types of data collected, reliability, validity, generalisability and ethics. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Laney et al. to explicitly support their point. These answers tended to score Level 4 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses, all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by Laney et al. as examples, which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. In addition, some candidates were following a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). Therefore, some candidates were producing prepared essays for Laney et al. without one of their points being about independent measures. A response that fails to have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did use independent measures in an evaluative sense but did not fully explain why it could be a strength and/or a weakness. Stronger responses could identify the potential strength of reducing aspects like demand characteristics. This then made that evaluative point 'in detail'. Several responses did not cover the named issue, only described what independent measures are or incorrectly evaluated the named issue by giving strengths/weaknesses of using repeated measures. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with example(s) from the study to show clear understanding. These are the requirements for a Level 4 response.

PSYCHOLOGY

Paper 9990/22
Paper 2 Research Methods

Key messages

- Paper 2 is a research methods paper asking candidates a range of questions, in this case including ones about the core studies in relation to aims, sampling, variables and ethical guidelines in relation to animals. In addition, more general questions asked about research methods, terms and concepts used to describe or evaluate research methodology, and application of this knowledge to both familiar and unfamiliar contexts. Such questions require different skills, some of which presented difficulties for many candidates. It is therefore essential that candidates are prepared for each of these skills, especially recalling concepts and the application of this knowledge.
- Applying knowledge and understanding to novel scenarios is important to success on this paper. This could have helped candidates in two ways:
  - Candidates needed to be able to apply research methods terms and concepts to scenarios presented in questions. These can include, for example, planning, criticising or developing designs or analysing data.
  - Candidates must take note of questions which indicate the need for a link. When a question says 'in this study', or makes direct reference to the scenario, responses must go beyond simply describing or evaluating, they must contextualise the answer in a relevant way. Candidates therefore need to be prepared for questions using this format and practice can help them to learn both how to extract relevant ideas and how to make novel suggestions based on scenarios.
- It is important therefore that candidates understand the basic research methods well and that they respond to the question by using the method stipulated by the question. Furthermore, to learn to identify flaws in a design (whether their own, as in **Question 10**, or one from a novel scenario for example in ***Section B***) also relies on having had experience of practical problems in conducting studies. This is a high-level skill and is most readily developed through practical work with designing and conducting small studies or through practice with novel scenarios. The overall format of **Question 10(a)**, and the nature of the mark scheme, is consistent between papers and years. Therefore, it may be helpful to prepare candidates by using past papers.

General comments

In general, candidates were able to access marks across the whole paper. However, very few were able to accurately and consistently demonstrate knowledge and understanding or to access the additional marks for linking their response to the scenarios, thus limiting performance as a whole. Nevertheless, some candidates demonstrated a good grasp of a range of psychological concepts and so were able to access the basic marks with these.

Candidates across the ability range were able to demonstrate some knowledge of a range of aspects of research methods in this paper. Success was greater on more straightforward questions such as **Questions 1(a)**, **1(b)**, **7(a)**, **8(c)(i)**, **8(c)(ii)**, **9(a)** and **9(b)**, than on more demanding ones, such as **Questions 8(b)(i)** and **8(b)(ii)** and **9(c)**. This examination tested a cross-section of psychological skills and on some, candidates showed limited success.

First, there were questions requiring accurate knowledge and understanding, including **Questions 1(a)**, **2**, **3**, **4**, **6** and **8(a)**. In such questions, covering for example concepts such as questionnaires (**Question 1**), aims and sampling (**Questions 2** and **3**) and participant variables (**Question 4**), candidates typically displayed some knowledge, such as that they were able to describe the meaning of questionnaires, aims and the independent variable, and were able to offer evaluation of some of these concepts, e.g. in **Question 1(b)**.

Second, there were questions requiring a link, for example to a study, which included **Questions 5**, **7(b)(i)** and **8(b)(i)** and **(ii)**, as well as **Question 10**. Here, candidates were sometimes able to earn partial marks for an initial identification of a relevant fact, such as describing volunteer sampling in **Question 3(a)**, but some were then unable to relate this information to this question. Similarly in **Question 8(a)**, some candidates were unable to link their knowledge of the independent variable to the scenario given and in **8(b)**, even fewer candidates were able to apply their knowledge of dependent variables to the scenario.

Question 10 was sometimes well quite answered although many responses were exclusively or partly based on other methods, such questionnaires. This meant that the candidate had used up valuable time unnecessarily and, as a consequence, the response lacked the necessary relevant detail to earn credit.

Comments on specific questions

Section A

Question 1

(a) This question part was generally well answered. Those candidates who did not gain the full two marks for their response generally did not indicate that in a questionnaire, participants must respond in written form (or online). Some of these candidates attempted to elaborate their (correct) description of a self report/asking participants questions with comments about open or closed questions. This, of itself, does not define a questionnaire.

(b) This question part was also generally well answered. Those candidates who did not gain the full two marks for their response generally did not offer a comparison (between questionnaires and interviews).

Question 2

(a) In response to this question, most candidates could describe an aim and many gave appropriate examples (e.g. from Andrade, although there were often good answers relating to Laney et al. and Baron-Cohen et al.). Although this question was very well answered, for candidates who did not score the full two marks, this was because they gave examples from outside the cognitive approach.

Question 3

(a) This question part was generally well answered but a significant number of candidates attempted to define volunteer with the word 'volunteer', i.e. repeated the question rather than answering it.

(b) This question required both knowledge and application. Some responses which did not earn credit suggested different sampling techniques, so were not answering the question. Others gave very general comments, such as '*include women*'. Nevertheless, there were many excellent answers.

Question 4

Although many candidates were able to give a definition of participant variables, far fewer were able to apply this concept. Both those candidates who were unable to go beyond 'participant variables are individual differences' and many of those who gave acceptable definitions (so understood what participant variables are) nevertheless cited variables such as gender or race. In this case, these were not participant variables but were independent variables within the study, so could not earn credit.

Question 5

Many responses here were generic, usually about sample size, with little reference to sleep and dreaming at all. However, some candidates provided excellent answers, such as ones referring to the possibility of differences in the dreaming of men and women.

Question 6

The guideline of 'numbers' was generally understood better than that of 'replacement', with many candidates being able to explain the concept and give examples from Pepperberg and/or Yamamoto et al. However, there were some irrelevant answers about studies of animals detecting numbers, about requiring greater numbers to avoid demand characteristics, about avoiding overcrowding or about keeping social animals in large numbers (this is 'species and strain'), the number of hours or studies they had been used for.

Replacement was less well understood with many candidates suggesting that this was about replacing distressed animals or ones that are in pain with new ones, replacing species with ones that will suffer less harm or replacing endangered species with other species. Such comments relate to the guidelines of 'species and strain' or 'pain and distress'.

Section B

Question 7

(a) Most candidates scored marks on this question. However, many responses did not focus on how to record data in a structured observation, i.e. candidates tended to produce partial answers. A small number of responses referred to collecting data using a bar chart, so scored no marks.

(b) (i) Many candidates earned at least one mark, with better application of the answers to the context than in **part (b)(ii)**. However, a significant minority of candidates confused a 'small number of species' with a small sample and others appeared to carry over from **Question 6**, suggesting that this would address the ethical issue of 'numbers'. The question however was about fewer *species* not fewer *animals*.

(ii) Most candidates earned one mark. Many answers limited to 'improve generalisability' without references to species, i.e. to what could be generalised, or to what other populations.

Question 8

(a) Many candidates were able to give a good definition of the independent variable but fewer were able to apply this to their answer: although many responses earned full marks, a significant number were incomplete, for example only providing one condition/level of the IV or just saying '*the drug/drugs*', without specifying that there were *different* drugs. A small number of candidates gave the IV as the stimulus or 'attention'.

(b) (i) Some candidates were able to suggest suitable measures, with some excellent inventive ideas, although many answers were vague, for example simply saying '*she could watch them*'. A minority of candidates scored no marks for this question part as they gave irrelevant answers, such as '*by using deception*' or suggested measures of emotional wellbeing and others were unable to earn credit as they described a technique, such as '*by questionnaires*' without operationalising the measure.

Better responses suggested actual measures of attention, some using examples such as from Andrade.

(ii) This question part was linked to **part 8(b)(i)**, and candidates who had not scored well in **part 8(b)(i)**, often gave confused or incorrect responses to **part 8(b)(ii)**.

(c) (i) Responses to **part (c)(i)** were generally good, although the responses were often generic, so restricted to one mark. A common confusion was between '*demand characteristics*' (the correct response to this question) and '*social desirability*' (an incorrect response to this question).

(ii) Some candidates did not respond to the question set. Instead of 'Suggesting …*why'*, such responses described *how* the participants were deceived (e.g. by being lied to/misinformed/not given full information).

The most common correct answer was that there was a lack of informed consent.

There were a small number of responses in which **parts (c)(i)** and **(c)(ii)** did not match.

(iii) The most common correct answer to **part (ii)** was commonly also correctly addressed here in **part (iii)** with the response 'debriefing'.

Question 9

(a) **Part (a)** of this question was generally well answered, although some responses attempted to argue that productivity or a similar measure could be used to indicate hours worked.

(b) P**art (b)** of this question was also generally well answered, although here too some responses attempted to argue that productivity or a similar measure could be used to indicate happiness.

(c) **Part (c)** of this question was less well answered than **(a)** or **(b)**, with most answers simply offering 'representative' or 'generalisability' without a link to why this would be important in this context, i.e. in terms of assessing hours worked or happiness. Nevertheless, there were some excellent answers noting that different jobs may have different working hours, tend to employ people with different personalities or offer working conditions that are more conducive to happiness.

Section C

Question 10

(a) **Part (a)** of this question produced a wide range of answers in terms of quality. There were some excellent responses, but very few achieved Level 3.

Many candidates gave little if any operationalisation of the independent variable, with some not even identifying the two levels given in the question. Similarly, few paid attention to the importance of the location, a critical part of their description for a field experiment. There was often no detail '*in a park*' although better candidates provided additional information such as '*when children normally play*', '*in areas that allow ball games*' or '*when adults are out jogging/coming home from work*' etc. There was, however, usually lots of detail of procedure often giving excellent information about how situations would be set up to enable the measurement of the dependent variable. Few candidates described controls, although better responses did so, e.g. '*all children trained away from the mall so they would all kick the ball into the fountain and look sad and helpless*'. Responses in the lower bands often appeared confused about the independent variable, suggesting this was the children, the children's clothing or the ball. Another source of confusion was the difference between a brief (at the beginning) and a debrief (at the end). Finally, a minority of candidates spent time presenting (fictitious) results, that would have been more usefully spent giving detail for the procedure.

(b) There were some good answers to **part (b)** of this question, typically focusing on a lack of control of extraneous variables. Only a very small minority of candidates ignored the rubric saying 'Do <u>not</u> refer to ethics or sampling in your answer'.

PSYCHOLOGY

Paper 9990/32
Paper 3 Specialist Options: Theory

Key messages

Questions 1(a), 3(a), 5(a) and 7(a)

It is important that candidates are made aware of the terminology, concepts and disorders identified in the syllabus as well as key terms used in named theories and studies as some were unable to identify and/or define the terms given in these type of questions. Creating a glossary of key terms, revision of terminology using flash cards and class quizzes on terminology could prove useful. Where the response gave an example to help define the term this often achieved full marks. These questions are worth 2 marks and a brief response is appropriate.

Questions 1(b), 3(b), 5(b) and 7(b)

These questions could ask the candidate to describe a theory, study, treatment or technique such as a self-report used by psychologists that is named in the syllabus. These questions could also ask the candidate to describe a part of one of the named studies, such as the procedure or the findings, or a summary of the key features of the study. This question is worth 4 marks and the candidates should write a more extended answer. It would be helpful for candidates to create a revision flashcard or mind map of each bullet point in the syllabus. The flashcard should be given the title used in the syllabus, for example, cognitive treatment and management of obsessive-compulsive disorder: Lovell et al., to help the candidate to identify which part of the syllabus the question is referring to. If the question asks for a part of the study, the response should only describe this part. For example, if the question asks for the procedure then the response should not describe the findings of the study.

Questions 1(c), 3(c), 5(c) and 7(c)

These questions could require the candidate to explain up to two strengths or weaknesses of what they have described in the **part (b)** of the question. The question could also ask the candidates to make a comparison or to evaluate using a specific issue such as the effectiveness of controls used in a study. This question is worth 6 marks so the candidate should write a more extended answer for each issue raised. Some responses were very detailed for one issue but then only briefly discussed the second issue. In addition, many of the responses were general and not specific to the study named in the question. To improve, responses should give specific examples to support their point. As mentioned for the odd question **part (b)**, the candidate should make a flashcard/revision notes and could include in this strengths and weaknesses of the theory, study or technique to help candidates prepare for these questions.

Questions 2(a), 4(a), 6(a) and 8(a)

These questions will always come from one of the bullet points in the syllabus. Candidates could describe the three or four studies, theories or techniques identified in the specification under the appropriate bullet point. For this exam, some of the answers used the incorrect topic area in the syllabus or the description was brief. It is possible for the responses to achieve full marks by describing at least two of the studies, theories or techniques and this would need to be a very detailed description. Full marks can also be achieved by a response that described three of the bullet points in detail (in less depth than if the response described two of the studies, theories or techniques) with excellent understanding and good use of terminology throughout. It is also important that the descriptions are linked to the topic area named in the syllabus. It could be useful for candidates to create revision notes with the title of each bullet point as the header in their notes.

Questions 2(b), 4(b), 6(b) and 8(b)

These questions will always ask the candidate to evaluate the theories, studies and/or techniques described in **part (a)** of the question. The response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. However, most responses that evaluated using two issues in this exam, achieved in the lower bands due to the response being superficial and often with little analysis. Some responses that considered three issues achieved higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the theories, studies and techniques described in the **part (a)** of the answer. The candidate must also provide some form of analysis. This could be done by discussing the strengths and weaknesses of the issue being considered, presenting a counter-argument to the issue under discussion or comparing the issue between two studies and/or theories. A conclusion at the end of each issue would be helpful in order to show excellent understanding of the issue under discussion. In order to achieve the requirements of the level 3 and 4 band descriptors it would be best if the response was structured by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure the answer covers this requirement of the question.

Some of the responses covered other issues rather than the one named in the question. Quite a few of the answers were structured by study/theory rather than by the issue which often led the response to be quite superficial and repetitive. A number of the responses did do analysis. Candidates should be aware these questions are worth 10 marks and attempt to include an appropriate amount of information.

General comments

The marks achieved by candidates for this series achieved across the full range of the mark band. Many candidates were well prepared for the exam and showed good knowledge, understanding and evaluation throughout their responses. Some candidates were not as well prepared and showed limited knowledge and understanding with brief and/or superficial responses. These candidates often had limited evaluation skills.

Time management for this paper was good for the majority of candidates and most attempted all questions that were required. A few candidates did not respond to one or more of the questions asked in the option area. A very small number of the candidates attempted to respond to more than two topic areas but often did not attempt all of the questions for each option chosen. These responses achieved at the lower end of the mark band.

The questions on abnormality were the more popular choice of option, followed by organisations.

Comments on specific questions

Psychology and Abnormality

Question 1

(a) Some responses addressed the question and achieved one or two marks by outlining common obsessions in obsessive-compulsive disorder (OCD). Popular responses includes obsession with perceived bodily flaws, fear of contamination leading to disease and obsession with hurting self or others. However, a significant number of responses outlined compulsions (washing or compulsive checking) rather than obsessions and some named types of obsessive-compulsive disorder, e.g. hoarding. These types of responses did not achieve any marks as they did not address the question asked.

(b) A significant number of responses were able to describe the Lovell et al. (2006) study on a cognitive treatment for OCD. Good responses were able to identify that the participants either took part in the face-to-face treatment or the telephone treatment. Most were able to describe the result that there was no significant difference in the effectiveness of both types of treatment. Some could identify the measures taken during the study (e.g. Y-BOCs), number of participants and number of sessions and length of these sessions. Weaker responses gave fewer details of the study with a brief outline of the treatment groups and results. Some responses described the incorrect study such as the cases studies by Lehmkuhl et al. or Rapoport. These types of responses were not creditworthy.

(c) The responses to this question covered the full range of the mark scheme. Stronger responses identified a strength and a weakness and explained this in terms of the Lovell et al. study with a clear example. Common strengths included the effectiveness of the treatments and why it is helpful to be able to use telephone treatment and standardised procedures so both had treatment over same period of time (10 weeks) allowing a direct comparison. A common weakness given was the reliance on self-reports and the consequent social desirability of participants' responses. Weaker responses had a lack of depth in their explanations and/or inaccuracies regarding the details of the study. An example of this was stating the study lacked generalisability because participants came from one place but failing to say what that place was and what the implications are for coming from one location to the generalisability of the research. These types of responses achieved in the lower mark band.

Question 2

(a) Responses varied considerably for this question and covered the full range of the marks available. Some responses highlighted how well prepared some of the candidates were for this exam whereas others showed very limited knowledge of this topic. There were many responses that were detailed, accurate and coherent with a good use of psychological terminology. The best responses focused on describing three treatments for impulse control disorders and non-substance addictive disorders in detail rather than attempting all four. There was good knowledge displayed of biochemical treatment and covert sensitisation. The main errors made in some responses were failing to describe the actual treatment given and relying instead on unnecessary detail of the case studies (e.g., how long the participant in the Glover study had had kleptomania rather than the fact that the treatment relies on the principles of classical conditioning). Very few responses outlined treatment and management of the other disorders in the syllabus.

(b) Many of the responses achieved in the level 1 or level 2 mark band with a few providing clear analysis and details of the treatments to back up their evaluative points, that enabled these type of responses to achieve level 3 and above. There was a tendency for responses to focus on many issues per treatment rather than applying the issue to the different treatment. These types of responses tended to identify that the treatments were deterministic or not, supported individual or situational explanations, etc. without any reason given for this or any analysis. A few responses did effectively discuss the named issue of quantitative data with some clear analysis. Many of the responses stated there was no quantitative data in any of the research which was incorrect. Apart from this named issue, common areas of discussion included individual and situational explanations, freewill and determinism and nature versus nurture. A few responses described treatments in **part (b)** which was not creditworthy here and should be put into the **part (a)** of their response.

Psychology and Consumer Behaviour

Question 3

(a) There were a few strong responses to this question with a clear explanation of the partially compensatory strategy in consumer decision-making. A common error made in responses was to describe the compensatory or non-compensatory strategy which was not creditworthy. In addition, responses that did not receive any credit were also where the response gave an anecdotal answer attempting to explain how companies try to convince consumers to buy products or what companies could do to get a consumer to purchase their product instead of their competitors.

(b) There were some very clear and detailed responses describing the Knutson et al. study on pre-cognitive decisions in consumer decision-making. These types of responses described the SHOP task in detail. Common descriptions of the procedure included the amount of money given to the participant, the timings in the procedure, how the products were selected and the number of products presented to the participants. Some responses did give the results of the study but this was not creditworthy. A significant number of responses either described the incorrect study or did not attempt the question.

(c) Where the candidate had been able to give some details of the study in **part (b)**, they were often able to identify controls used by Knutson et al. but few were able to give detail of how effective they were. Stronger responses focused on a couple of controls in some detail usually the timings or the exclusion of those who had made too many head movements. Weaker responses listed some of the controls used in the study without any attempt made to explain the effectiveness of them. If the candidate did not know the study in **part (b)** of the question, **part (c)** was frequently not attempted, a general evaluation given or a guess was made at a control that might be used in a lab study. These types of responses usually did not receive any credit.

Question 4

(a) Many responses achieved level 2 or 3 for this question. Good descriptions were often seen in responses for the Robson et al. study on space at restaurant tables, although there were a significant number of responses that stated the study took place in a restaurant which was incorrect. There were also some good descriptions of the Milgram et al. study on defending place in a queue. Many of the responses did describe one or more of the theories of personal space with overload being the most common. However, the vast majority of these responses did not provide any link to consumer behaviour. Weaker responses gave very brief or inaccurate details of the research or just described types of personal space without any reference to theories of personal space or the research. These types of responses achieved in the level 1 mark band.

(b) Those responses that achieved in the higher mark bands for **part (a)**, tended to produce good answers to this question, with some understanding of how generalisability applies to the studies with some examples given. Some attempted analysis, for example pointing out that a sample was large (Robson et al.), which would increase generalisability but that they were from the same area geographically and thus lowering generalisability. Ethics and applications to everyday life were other commonly chosen issues. Many responses used inappropriate evaluation issues such as nature versus nurture, determinism versus free will or individual versus situational factors. These types of responses struggled to be relevant to the research and theories they had described in **part (a)**.

Psychology and Health

Question 5

(a) This was often well answered with many responses achieving full marks with good explanation of hypochondriasis. Many responses described a fear of illness, regularly going to the doctor and not believing the diagnosis, and examples such as believing a mild headache was a brain tumour. Some responses confused hypochondriasis with Munchausen syndrome which was not creditworthy.

(b) There were some strong responses to this question with some giving a detailed description of the Savage and Armstrong study on practitioner consulting style and patient satisfaction. Common details of the study included a description of the sample, how the allocation to conditions was done in addition to the 'directing' and 'sharing' styles and the results that the 'directing' style was preferred. Weaker responses often just identified the two styles and gave a result from the study. It was common error for responses to describe the Bryne and Long study rather than the Savage and Armstrong study or to state that the 'sharing' style was the preferred style.

(c) Most responses identified at least one, if not two strengths of the Savage and Armstrong study. Stronger responses focused on the random allocation to conditions and the usefulness of the results. Most responses achieved level 1 or level 2 due to not providing any clear example from the Savage and Armstrong study to back up their strength. Responses that were not creditworthy typically evaluated the incorrect study.

Question 6

(a) The responses to this question covered the full range of the mark scheme. Stronger responses had a good understanding of the measures of stress and were able to either give clear details of the recording devices and sample tests as well as one of the self-report questionnaires with the Holmes and Rahe stress scale being the most common. Some responses described the research that investigated the measures of stress in some depth. Weaker responses lacked details and were sometimes confused about how an fMRI measures stress or gave incorrect details of the research.

A few responses were anecdotal and described measurements of stress such as rating scales and heart rate with no details of these measures or any reference to research. These types of responses typically achieved in level 1.

(b) There were some good responses to this question. These were often able to evaluate the named issue of validity of the measurements that they had described in **part (a)**. A few attempted some analysis and were able to discuss how a measurement could provide a valid measure of stress while also identifying the problems with the validity of this measure or a comparison was made between the validity of the physiological and psychological measures. Other common evaluation issues for this question included the reliability of the measurements and/or research, self-reports (although this sometimes just repeated the issues of reliability and/or validity), quantitative data and objectivity versus subjectivity. Weaker responses seemed to be when a large number of issues were covered but all were done in a very superficial manner, simply stating that issues did or did not apply to studies and/or measurements but not why.

Psychology and Organisations

Question 7

(a) There were many good responses to this question with many achieving full marks. Common physical work conditions included the layout of the workplace (e.g. open plan) and lighting. Common psychological conditions included bullying, recognition and satisfaction. A significant minority of responses gave two physical work conditions or a great number of work conditions including a mixture of physical and psychological.

(b) There were a number of good, detailed responses to this question. Some responses described the safety promotion campaign by Cowpe in depth including the different strategies, the 10 TV regions and how the results were obtained. Those who attempted this question often included the actual findings (which were not asked for) and this part of the response was not creditworthy. A significant minority of responses confused this study with Fox et al. on token economies or gave anecdotal responses about safety in the workplace.

(c) The responses to this question covered the full range of the marks available. Stronger responses made reference to high population validity or effectiveness as strengths of the safety promotion campaign. The most common weakness cited was the difficulty of knowing if the reduction in fires was due to the advertisements or not as data was only collected from the fire brigade. Weaker responses lacked detail and did not include any clear examples from the Cowpe campaign to back up their strength and/or weakness. Those candidates that described the incorrect study in **part (a)** did not usually achieve any marks for this question or achieved level 1 due to briefly outlining the effectiveness or good population validity of the campaign.

Question 8

(a) There were some good, detailed responses to this question, with clear details given of workplace sabotage (Giacalone and Rosenfeld, 1987), absenteeism (Blau and Boal, 1987) and measuring organisational commitment (Mowday et al., 1979). The Giacalone and Rosenfield and the Blau and Boal study tended to be the most detailed of these descriptions. Weaker responses tended to be brief or gave anecdotal descriptions of attitudes to work. A number of responses referred to other sections of the syllabus such as motivation to work and bullying at work. These types of responses were not creditworthy.

(b) There were some good responses to this question where it was structured by evaluation issue and began with the named issue of practical applications. Some of these responses gave clear examples from **part (a)** of their response to back up their evaluative points and attempted some analysis. In addition to practical applications, popular evaluation issues were generalisability, cultural bias and validity. Weaker responses did attempt to evaluate the named issue but did not explain why the research had a practical application and did no analysis. These types of responses tended to state that a theory or study either was or was not applicable to companies with no explanation of this given. Those responses that attempted individual and situational explanations were often very weak with a lot of confusion over the definition of this issue.

PSYCHOLOGY

Paper 9990/42
Paper 4 Specialist Options: Application

Key messages

- What has been learned from the AS component of the syllabus should be transferred to the A2 component. For example, at AS candidates learn about methodology, such as experiments, which also apply to A2.
- Questions should be read carefully ensuring that the focus is on what the question asks.
- All components of the question should be included in answers. For example, part (d) for **Questions 1**, **2**, **3** and **4** required advantages and disadvantages (plurals) *and* a conclusion.
- In ***Section B***, **Questions 5**, **6**, **7** and **8**, methodological knowledge must be evident and detailed for top marks to be accessed. The procedure, however detailed is just one methodological aspect. For top marks, answers must explain methodology rather than merely identify it.
- In ***Section C***, **Questions 9**, **10**, **11** and **12**, to access top marks answers must include a debate which has two sides, such as strengths/advantages and weaknesses/disadvantages. Supporting evidence should also be provided.
- Psychological knowledge should be applied wherever possible. Anecdotal and common-sense answers will not achieve top marks.

General comments

Section A

- Candidates frequently failed to address the 'stem' of the question, the introduction or the opening words in ***Section A*** when this is crucial to answering each question part that follows.
- Answers must refer to the study the question is about. Many answers made general comments rather than focus on the study itself (see specific questions below for examples).
- Many answers correctly included advantages/strengths and disadvantages/weaknesses but many answers did not relate these to the question and so restricting marks. For example, to score one mark, answers must include an advantage and this must be related to the question.
- Many conclusions repeated what had already been written, and such summaries scored no marks. A conclusion is a 'decision reached by reasoning' and so as the reasoning has been done through the advantages and disadvantages, a final decision/conclusion needs to be drawn.
- Candidates should think about what the question requires rather than writing pre-prepared answers. Many questions will test the ability to apply knowledge from one thing to another, particularly methodological knowledge.
- Candidates should always provide sufficient detail to score all the available marks. A single sentence is more likely to score one mark rather than two marks, so a little elaboration, explanation or example that goes beyond the basic sentence is always recommended. Candidates should always try to impress the Examiner with their psychological knowledge.

Section B

Answers to **part (a)** questions in this section should include an appropriate design, have applied a range (four or five) of relevant methodological design features, each of which should be explained fully, showing good understanding. Many answers listed features such as 'I would have a random sample' and 'It would be an independent measures design' without explanation of why it would be a random sample, or how this would be obtained.

In **part (b)** answers should explain the methodological decisions on which their **part (a)** design is based and also explain the psychological evidence on which their design is based. Merely describing a relevant piece of research from the topic area is insufficient. The links between the research and how it informed the design must be shown. Further, there is no need for a name (date) to be quoted for each sentence, with some candidates writing 'I chose a self-selecting sample because Milgram (1963) did' for example. This just identifies a study using that technique. It does not explain the choice of sampling technique.

Section C

It is essential that answers focus on the question that is set. Every question in this section invites candidates to consider the extent to which they agree or disagree with the statement, rather than to describe everything they know about that topic area, and answers that fail to address the question will only achieve minimal marks. To score marks at the top end of the mark range, answers must focus on arguments both for and against the statement, answers must the use appropriate evidence to support the argument, and at the very top of the mark range, answers should show awareness of wider issues and evidence that is relevant.

Comments on specific questions

Section A

Question 1

(a) To score full marks, candidates were required to give two conclusions. Most candidates provided two different conclusions taken directly from Fig 1.1 and were awarded maximum marks. Some candidates gave only one and some candidates repeated the same conclusion twice.

(b) Many candidates scored full marks by writing about two of the three ways used by Ost and Westling which were: (i) independent assessor ratings (by a qualified therapist) (ii) self-report scales (providing quantitative data) (iii) self-observation of panic attacks (keeping a diary of the frequency and severity of attacks). Some candidates appeared to guess at questionnaire, interview and observation, all of which were awarded no marks.

(c) (i) Many candidates scored full marks for explaining the difference between applied relaxation (AR) and cognitive behaviour therapy (CBT). Often the simple point was made that applied relaxation is tensing muscles done by the patient without talking, whereas CBT is a 'talking' therapy.

(ii) The essential difference is that progressive muscle relaxation decreases blood pressure to counteract the effects of stress-related hormones whilst applied tension (to muscles) increases blood pressure to prevent the person from fainting. Many candidates wrote about this difference and scored full marks. Many candidates did not focus on physiological effects and did not answer the question set.

(d) At the top end of the mark range candidates included two strengths plus two examples, two weaknesses plus two examples and a conclusion. Other candidates included two strengths and two weaknesses, but only scored partial marks because anxiety disorders, as the question required, were not mentioned. A few candidates did not provide strengths and weaknesses and described a number of behavioural techniques. These answers scored no marks.

Question 2

(a) Many answers scored one mark for stating that a dependent variable (DV) is what is measured and in addition most candidates scored the second available mark by elaborating (e.g. the IV is the cause and the DV is the effect) or by giving an example (in the study by Guéguen et al. the DV was the average amount of sales for each customer).

(b) Many candidates scored partial marks in response to this question. The crucial component of any dependent variable (DV) is the actual measurement because without knowing this, no DV is operationalised. For example, many candidates correctly identified 'sales rates' but did not state that this was measured in 0, 1 or more items. Similarly the 'length of stay at the stall' was identified without stating that this was timed in minutes. Full marks were awarded to candidates including fully operationalised DVs in their answers.

(c) Most candidates scored two marks out of four for providing two advantages of qualitative data. However, the question included the crucial words in this study which many candidates did not address. Without this the advantages could apply to any study and the required skill for full marks is to be able to apply the advantage to the study in question. For example, writing 'a reason can be provided by participants' (scored one mark for the advantage) for why they spent more money at the stall (scores a second mark because this is what happened in the Guéguen et al. study).

(d) To score full marks advantages and disadvantages must be related to the study or topic in question. Many answers made no mention of background music or consumer behaviour; answers were merely a list of advantages and disadvantages of field experiments.

Question 3

(a) Many answers scored one mark for providing a basic statement of unrealistic optimism, for example 'it is where people tend to think they are invulnerable' and in addition most candidates scored the second available mark with some elaboration or by providing an example such as 'people think they will not have a heart attack before age 40'.

(b) (i) This question required an outline of the measures of optimism questionnaire used by Weinstein. The questionnaire (one mark was awarded for any two of these features) included 42 events, 18 positive events, 24 negative events, positive and negative being intermixed. Credit was also awarded for any other correct feature. Any of these features showed that the candidate had knowledge of the Weinstein study. A number of candidates were not able to demonstrate knowledge of what the questionnaire might have included, most of these answers scoring 0 marks.

(ii) The scale used by Weinstein had choices ranging from 100 per cent less than average, through to 0 per cent average, to 100 per cent more than average. Stating these three aspects would have scored the full two marks. Again, a number of candidates were not able to demonstrate knowledge of the scale.

(c) Partial marks were often scored in response to this question. A way to assess unrealistic optimism was often correctly identified, such as using open-ended questions, an interview or an observation, and these responses would be awarded one mark. However, the suggestion needed to be related to unrealistic optimism to score full marks. For example, writing 'I would use an interview' would score one mark, with 'why do you think you will not have a heart attack before age of 40?' would score one further mark.

(d) This question asked for advantages and disadvantages of assessing unrealistic optimism with students as in the Weinstein study, although some responses focussed on young children. Many good answers provided both advantages and disadvantages, but only a few candidates went on to support each advantage/disadvantage with an example. Many candidates summarised their answer at the end, which scored no marks. The question requires a conclusion, a 'decision reached by reasoning', not a summary of what has already been written.

Question 4

(a) Some candidates scored full marks and others a partial mark, but there were also many who were unable to demonstrate their understanding. A full (two mark) explanation of affiliation might be: the need to be liked and accepted by other people; effort is applied to creating and maintaining social relationships and friendships at work.

(b) This question part required two reasons why affiliation is important in a work environment. Marks were awarded across the full range (0 to 4). Answers scoring full marks included: In order to be effective, the workers in a team should like each other removing any group conflict; people are social and need to feel respected by others, to be treated fairly and not bullied.

(c) The needs of a manager are likely to be different from the needs of a worker and many candidates understood this, explaining that the manager would have a need for power (being a manager) and for achievement, with affiliation being less likely. Workers would be more likely to have a need for affiliation and achievement with power being less likely. Some candidates focussed instead on Maslow's need theory.

(d) This question required a discussion of using self-reports to measure need for achievement and very few answers referred to need for achievement.

Section B

Question 5

(a) Nearly all candidates chose to write about an experiment and most included a range of relevant features. Answers scoring high marks showed a good understanding of IV and DV, whereas others failed to score marks through a lack of explanation. For example, many candidates stated that the IV was age. However, this could not be operationalised and inclusion of specific age ranges was needed, such as 20–30, 30–40, 40–50. General methodological features were also included such as the sampling technique and although most answers were appropriate, there was a lack of coherence in others. For example, many candidates decided to advertise 'online' (worldwide? how?) for a sample to participate in a laboratory experiment (in a specific country?) without an explanation of how this would work. More time spent explaining features of the plan would result in more marks than providing a long list of unexplained features. List like answers such as 'IV = age, sample = opportunity sample, DV = effectiveness' score very few marks.

(b) Many candidates for their psychological evidence correctly wrote about rational emotive therapy and Ellis's A, B, C, D and E, although only rarely explained how these related to the plan of their study. If effectiveness (of REBT) is being measured then there should be some way in which this effectiveness can be assessed, even if it is as simple as no longer visiting the REBT therapist. Methodologically experimental designs were good, but often lacking explanation. For example, choice of DV was rarely explained. Choice of experimental design was stated, but with no explanation of why an independent design had to be used. More explanation showing understanding is the key to scoring high marks.

Question 6

(a) If the method to be used is a questionnaire there is no need to apply the features of an experiment, which is what most candidates decided to do. Writing about features such as IV and DV, etc., were unnecessary here and not a good use of time. More important were the details of the questionnaire such as the type of question (open, closed or a combination), setting (where the questionnaire will be conducted), example questions, how the answers would be scored, and how they would be analysed. In addition to these specific features, candidates should also write about general features such as the sample and sampling technique, type of data (quantitative or qualitative or both) and possible ethical issues. In this question candidates could choose to focus on one feature, lighting, colour or smell, or, given that the method was a questionnaire, they could include all three.

(b) Psychological knowledge in the stronger answers showed a good understanding and application of a range of relevant studies, such as those by Kutlu et al. (2013) on lighting and Chebat and Michon (2003) on odour. A few candidates referred to music and the study by Guéguen et al. (2007) which was allowed because although it was not mentioned in the stem the question stated: 'variables *such as*…' and music is an atmospheric variable. Methodological explanations for the choice of questionnaire type, setting and example questions were lacking.

Question 7

(a) This question required candidates to use an interview. Many candidates decided to make the study an experiment and wrote about IV, DV, design, etc. These features are not part of an interview. Inclusion of specific features of an interview do score marks and include features such as the interview technique (face-to-face or telephone), the interview type (structured, semi-structured or unstructured) the type of questions asked (open, closed or a combination), setting (where the interview be conducted), example questions, how the answers would be scored, and how they would be analysed. All these features, along with general features such as sample and sampling technique, needed to be related to helping improve patient adherence.

(b) The focus of this question was on improving adherence and so the work of Ley (1988) would have been an important inclusion. Instead, many candidates focused on aspects of the patient-practitioner relationship (a different syllabus topic area). Whilst this was not irrelevant and some marks were awarded, the work of Ley had to be included to score full marks. Methodological reasons for the choice of interview features were often lacking mainly because candidates had not included the features in **part (a)**.

Question 8

(a) Many candidates decided to make the method an experiment. Whilst many experiments do use observations to gather data, a 'stand-alone' observational study can be conducted with no reference at all to an experiment. Many candidates knew and applied the specific features of an observation, although some were inconsistent in applying all the features. There are four essential elements that make up an observation: setting (controlled or naturalistic), participants (overt or covert), observers (are they participant or non-participant) and data (structured or unstructured). Some candidates, often those scoring high marks, included event or time sampling and how inter-observer reliability could be assessed.

(b) The psychological evidence included in this question part should focus specifically on explaining the basis of the **part (a)** plan. For this question, candidates described Tuckman's stages, forming, norming, storming and performing, and scored marks, but very few explained how these stages related to what was observed in **part (a)**. Methodological knowledge was often very good, with candidates explaining why they had used a covert rather than an overt observation for example.

Section C

Question 9

Many candidates did not score high marks because they did not focus on the question set. Many candidates appeared to write all they knew about genetic and alternative explanations of phobias and answers resembled the description required in a Paper 3 essay. Such answers scored very few marks. This Paper 4 question is very different: answers should be organised around the question 'to what extent' with a debate presenting reasons and evidence explaining why the statement is agreed with or disagreed with.

Very few candidates answered this question in this way; most described the genetic explanation followed by a description of the cognitive, behavioural and psychoanalytic explanations.

Question 10

Three types of answer were evident. There were answers which answered the question, structured the answer and give reasons why they agreed with the statement followed with arguments why they disagreed with the statement, scoring high marks. Many candidates described the Snyder and De Bono study on consumer personality, but this is not a 'describe' question, it requires a discussion of consumer personality and individual differences. Some candidates described and then discussed consumer personality but needed to give an opinion about the extent to which they agreed with the statement.

Question 11

Most answers in response to this question were weak, with very few candidates writing about postal questionnaires. Postal questionnaires were used in the study by Yokley and Glenwick and in the study by Riekert and Drotar. The question stated 'assess improvements in adherence' so the focus should have been on improving adherence, but many candidates described different ways in which adherence could be measured.

Question 12

Like other answers for different options in this section many candidates described what they knew, rather than organising and using information to address the question. For this question there were many anecdotal answers describing recognition, respect and a sense of belonging and, although these answers scored some marks, there was a lack of awareness that these are intrinsic motivators. The contrast with extrinsic motivators could have provided good discussion, but this was also absent from most answers.