PSYCHOLOGY

Paper 9990/12

Paper 1 Approaches, Issues and Debates

Key messages

Candidates need to know all components of every core study as listed in the syllabus. Questions can be asked about any part of a core 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, a named issue to be included or relate back to a previous answer. 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 with an explicit example from that study.

Candidates need to be able to know about 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. Again, this must be explicitly made by the candidate.

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. This can sometimes mean a candidate may run out of time for other questions.

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. This is a crucial skill to develop as some candidates appear to have good knowledge of a study but do not apply this effectively to the question(s) set.

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) The majority of responses correctly stated the other time interval used in the study by Dement and Kleitman.

(b) The majority of responses correctly named the apparatus used to measure brain activity. Incorrect responses tended to name brain scanning techniques like fMRI and MRI.

(c) There were several correct responses to this question. The most popular responses focused on the tomato throwing dream but there were other responses about driving a car. Incorrect responses tended to focus on dreams that had been reported in the study like climbing a ladder, but these would have shown *vertical* eye movement. It is important that candidates read questions carefully to ensure they are presenting the correct information.

Question 2

(a) There were few correct responses to this question. Popular correct items included fruit, nuts, and wood. However, the majority of responses did not appear to focus on the demands of the question. For example, there were many responses that focused on what Alex would verbally 'say' in trials (like shape) or on items that Alex did not have to request. Information within a journal paper that is reported by the author as part of the whole procedure needs to be known. There were a significant number of blank responses.

(b) There were some clear, concise responses to this question. For example, some noted that many of the test procedures were standardised meaning that it would be relatively easy to replicate. Weaker responses either provided one strength not related to reliability or chose to ignore the demands of the question and write about validity (the choice of test items was mentioned many times in relation to validity). It is important that candidates know the difference between reliability and validity.

Question 3

(a) There were many correct responses to this question. Popular descriptions included the sample size, where the children were enrolled and their age range. However, there were a significant minority of responses that did not describe three features/characteristics of the sample or confused the sample with that from the study by Baron-Cohen et al. It is important for candidates to note the number of marks assigned to a question as this typically corresponds (in short answer questions) to the number of correct elements that need to feature in a response.

(b) The majority of responses could describe a result that involved aggressive gun play. The most popular choice was boys showing higher levels compared to girls. With results, where possible, there must be a meaningful comparison to be able to score the marks available. Therefore, simply stating 'boys showed more' could only be awarded partial credit. Several incorrect responses chose a different result or results about mallet aggression.

Question 4

(a) Stronger responses could clearly describe four features of the shock generator and were awarded maximum marks. Popular features included the labelling of the buttons, the number of buttons, the voltage range and that each button went up by 15 v. Incorrect responses tended to focus on what happened once a button was pressed in terms of the reactions (or non-reactions) of Mr. Wallace. Some responses evaluated the usefulness of the shock generator, or outlined another part of the procedure in the study by Milgram.

(b) For this type of question, responses must contain two parts. The first is the 'what' – what real-world application could be. The second is the 'how' – how will the real-world application be achieved in an ethical way. Common responses included utilising uniforms or some form of verbal prods to train military or to keep children more obedient to instructions in the classroom.

Question 5

(a) The majority of responses could outline the aim of the study by Andrade. Popular choices included the role of doodling in memory and concentration. There were some responses that only gave a very brief aim which could only be awarded partial credit. Weaker responses presented a finding from the study which could not be credited.

(b) Stronger responses could clearly explain one reason for standardisation. The most popular explanation focused on reliability and replicability with an example from the study by Andrade. Some responses were awarded partial credit for explaining a reason but not addressing the 'in this study' part of the question. Some responses focused on control and cause and effect and gained appropriate credit. The question was why and not how so responses that simply provided examples of how the study was standardised could not be awarded credit.

Question 6

Stronger responses could clearly outline the individual-situational debate and provide clear examples from the study by Schachter and Singer. Popular examples included differences in reactions to the injection (individual) and how participants were affected by the behaviour of the stooge (situational). There were a significant number of responses that were tautological and could not access marks. For example, stating that the situational side of the debate is about the situation cannot be credited as it is simply re-using the words in the question and does not explicitly show understanding. To improve, candidates need to have examples from each Core Study that appropriately support each of the issues and debates at AS-Level. There was a significant number of blank responses.

Question 7

Strong responses clearly described the procedure between the two time points stated in the question. Popular elements included the use of photographs, having to rate the pictures on various scales, and that the participants were debriefed. Some responses provided information about a different part of the procedure, or different aspect of the study, for example, the sample or a result. Candidates need to know the procedure for all Core Studies. There were a significant number of blank responses.

Question 8

(a) Stronger responses outlined what was expected of participants in the recognition task. A significant number of responses misinterpreted the question and provided information about the emotional ratings given by the participants in the first part of the study. The recognition test is the second part of the study. It is important for candidates to be familiar with the exact procedures for all Core Studies.

(b) There were some strong responses provided here that fully engaged with the stimulus material. There was a roughly equal choice between supporting Govinda or supporting Ansh. Popular arguments supporting Govinda included the use of controls to establish some internal validity, and the use of objective measures. Popular arguments supporting Ansh tended to focus on the lack of mundane realism and low population validity. Some responses simply described aspects of validity without engaging in the Govinda-Ansh debate so could only gain partial credit. To improve, responses must clearly choose one side of the argument and then explain why they support it using examples from the study.

Question 9

(a) For these types of questions, responses should focus on the general psychology that is being investigated in the study rather than a specific aim of the study. Therefore, aspects of the Yamamoto study that could gain credit here included altruism, empathy, and prosocial behaviour. Credit could be given to generic descriptions of the principles of either of these. Weaker responses focused too narrowly on the aims of Yamamoto and what was achieved by the chimpanzee participants.

(b) The responses to this question varied a lot. Strong responses clearly showed understanding of *housing* using an example from the study by Yamamoto. However, many responses argued about it being cruel to not be in the wild. It is important for candidates to know the background of participants, especially when it is non-humans. *Numbers* was the guideline with the strongest set of responses, as candidates could clearly outline what the guideline was and how Yamamoto only used a few chimpanzees in the study. For *rewards,* responses needed to give both a description of what the reward system as well as link to the ethical guideline. It also needed to be in the context of using reward and/or depriving a chimpanzee to gain credit. The guideline of *species* was known by a minority of responses as many tended to simply describe the species used in the study, chimpanzees. To improve, answers must show explicit knowledge of an ethical guideline and then provide an example from the study that shows it was broken or not broken.

Question 10

The strongest responses evaluated the study by Saavedra and Silverman in depth and in terms of two strengths and two weaknesses, with at least one of these points covering the named issue of self-reports. Common choices included types of data collected, reliability, validity, generalisability, and ethics. Strong responses explained why an element of the study was a strength or a weakness, using specific examples from the study by Saavedra and Silverman to explicitly support their point. 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 Saavedra and Silverman as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough and well argued, with a fourth point that was brief which meant the response did not reach the top band in the main. Description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. In addition, some responses appeared to be following a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). Therefore, some responses appeared to be prepared essays for Saavedra and Silverman without one of their points being about self-reports. Responses lacking an evaluation point about the named issue cannot score top marks. Lower scoring responses briefly outlined strengths and weaknesses, giving only some or no context. In addition, many responses did use self-reports in an evaluative sense but did not fully explain why it could be a strength and/or a weakness. Several responses did not cover the named issue, only describing what self-reports were used in the study by Saavedra and Silverman. 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 examples from the study to show clear understanding.

PSYCHOLOGY

Paper 9990/22
Paper 2 Research Methods

Key messages

- This research methods paper asks candidates to answer a range of questions, including ones about the core studies in relation to research methods, terms and concepts used to describe or evaluate research methodology, and application of this knowledge to both familiar and unfamiliar contexts. Responses to this paper demonstrated a range of ability in these, although application of knowledge to unfamiliar contexts in short answer questions presented a challenge for many candidates. This is an essential skill and candidates benefit from the opportunity to learn to respond to scenario-based questions, particularly in relation to two skills:
  - Firstly, candidates need to be able to extract information from scenarios, enabling them to understand the situation or study being described. For example, to identify the variables, methods and techniques used in a study or the role of individuals in a social setting.
  - Secondly, candidates must be able to go beyond answers that simply identify, describe, or explain concepts, so that they can give answers linked directly to the content of the scenario described, for instance in response to question parts ending '…used in this study.'
- In addition, candidates should read and follow the instructions given in the question carefully. For example, recognising the difference in demand of questions asking them to 'state', 'describe' or 'explain'.

General comments

Candidates were able to access marks across the whole paper. Although the issues identified above tended to limit performance.

Candidates across the ability range were able to demonstrate knowledge of a range of aspects of research methods in this paper. Success was greater on some straightforward questions such as **2(b)** (null hypothesis), **7(a)(i)** and **7(a)(ii)** (behaviours that could or could not happen in a dream), **7(c)** (human ethics), and **8(b)** (suggesting novel factors to investigate).Most candidates were also able to score some marks on more complex questions such as **Questions 3(a)** (advantage of a single participant), **7(d)** (using eye movements), **8(c)** (drawing conclusions from graphs), **8(d)**(problems in observations) and **9(b)** (measuring speed of recall). However, there were also some questions which candidates found challenging, such as **4** and **9(a)**.

Where questions required identifying or explaining a term or concept, weaker responses often gave tautologous answers. For example, repeating a concept from the question in their answer (**Question 1(a)**) or attempting a description or explanation using a term itself (e.g. **Questions 2(a)** and **6**), which could not be credited.

Question 10 was sometimes very well answered. Many other responses were at least partly irrelevant, describing aspects of experimental designs. Candidates need to be able to make a clear distinction between open and closed questions – an essential part of a self-report such as a questionnaire. Weaker scripts gave inappropriate, incomplete, or inconsistent examples. Such responses therefore lacked the necessary relevant detail to reach full marks. To ensure that candidates can provide the necessary, relevant detail for each research method it is important therefore that they have practised this high-level skill. It is most readily developed through practical work with designing and conducting small studies in class or through practice with novel scenarios.

Finally, candidates should be discouraged from writing in pencil then writing over the top in pen, as this makes scripts almost illegible.

Comments on specific questions

Section A

Question 1

(a) Many responses stated one feature correctly, commonly the concept of an artificial environment, the other was often a repetition of the feature of controls identified in the question. In addition, a common type of incorrect response was to focus on strengths or limitations rather than features.

(b)(I, ii) Many candidates were able to provide one or two controls in response to **part (b)(i)** of the question, such as the use of a visually dull or quiet room. Few were able to answer question **parts (b)(i)** and **(b)(ii)** fully.

Most responses candidates were able to identify a feature that acted as a control although sometimes in a general way, such as saying that one control was the monotonous message. However, in **part (b)(ii)**, such candidates often explained what would have happened if the message had not been boring rather than differences in the responses in participants between conditions if the message had not been the same for everyone.

There was some confusion with control groups, with answers suggesting the difference in paper given to the two groups as a control, rather than recognising this as the IV.

Question 2

(a) This question part appeared to be found challenging. Strong responses need to explain their points clearly. Weaker ones often simply repeated the word 'direction'. In addition, many responses made reference to "increases or decreases" rather than "more or less" when regarding which gender would be more aggressive. There were also a minority of scripts in which the response given to this question was about the Bandura study, for example referring to children and imitation. Weaker replies gave irrelevant answers referring to correlational hypotheses.

(b) There were many correct responses here, although some offered correlational hypotheses. Another common mistake was to refer to differences between gender and aggression rather than gender differences in aggression.

Question 3

(a) The majority of responses earned marks here, typically in relation to being able to study in depth or the ethical guideline of numbers. Weaker replies usually simply referred to being able to study the participant in depth, without linking their answer to Pepperberg's study. A minority of answers incorrectly referred to collecting more data, rather than collecting data in more depth.

(b) There were many one-mark answers to this question part, with responses simply stating a lack of generalisability. Stronger answers referred to specific differences between Alex and other parrots, including intelligence and the time spent in captivity.

Question 4

Answers to this question tended to be either very good, sometimes including four or five creditworthy points, or were not able to show understanding. Some confused normal distribution with standard deviation and some referred to distribution of characteristics in a sample.

Question 5

This question was not answered well. Many responses simply referred to bar charts as presenting information clearly and simply without identifying that they are appropriate to use for categorical data. However the answers that did identify this almost always achieved full marks.

Question 6

There were some very good answers here, but many responses included evidence of confusion about these sampling techniques. Random sampling was often described using the word random repeatedly, and examples were often of opportunity samples, such as suggesting that a sample taken on the subway was random because random people were there. There is also some confusion between opportunity samples and self-selected samples, with responses suggesting the opportunity samples were where participants were offered the opportunity to participate.

Section B

Question 7

(a) (i, ii) Many responses used examples from the Dement and Kleitman study such as basketball or climbing a ladder. Flying was a popular answer for the second part of this question. Weaker answers did not identify behaviours, such as emotions or thoughts, and sometimes only suggested questions that could be asked about the behaviours in the dream.

(b) The most common answers here were to suggest using questionnaires rather than interviews, or the use of filler questions to disguise the aim. A minority of candidates explained how lying could be detected without referring to how it might be reduced.

(c) This question part was well answered with most responses accurately identifying privacy, with a small number focusing on protection from harm. Many demonstrated that they understood privacy fully and were able to distinguish this from a more generic response about confidentiality (which would not have earned credit). A smaller number of answers incorrectly identified confidentiality. Several weaker responses identified privacy correctly but did not give enough detail about this to achieve both marks.

(d) Many responses explained that the direction of eye movements could be used to determine whether the content of the dream reported match these eye movements. Some lacked relevance in only referring to REM and NREM and made no reference to the way that specific eye movements could be related to dream content.

Question 8

(a) Roughly as many scripts suggested this was a participant observation as suggested it was a non-participant observation. The correct explanation of either could earn full credit. Weaker responses simply identified the type of information with no explanation, which was not an explanation so was not creditworthy. Another common error was to offer explanations relating to covert or overt observations.

(b) Some good answers were given to this question, such as whether people were alone or in groups, their race, their dress, the use of headphones, their gender or whether they had a child with them. A minority of candidates suggested factors such as personality which were not observable. A small number repeated information about animals from the question.

(c) Most answered this well, although there were some who did not complete the comparison in their answer, simply saying that females were more likely to say hello and needed to include ‘than males’.

(d) Many responses to this question were generic, simply identifying demand characteristics or an ethical issue without applying this to the specific study. Candidates need to be able to distinguish between demand characteristics and social desirability.

Question 9

(a) This question part was not well answered with many responses identifying Tara’s study as an experiment rather than a correlation, and any such responses described the study as having an independent measures design. Even when responses correctly identified a correlation, they often did not give the two variables in Tara’s study, or simply said ‘because she is looking for a relationship between two variables’, so had not linked their answer to Tara’s method.

(b) Correct responses were frequent with responses saying that they would show participants a list of words and then test them. However, there was often no indication of timing. Others offered a number of correct answers rather than time taken to recall the answers. Some weaker responses did not display an understanding of the correlational nature of the study, for example not recognising that the variables were independent from each other.

(c) A very large number of responses referred to participant variables here, showing a misunderstanding of the study. A smaller number correctly identified problems of timing accurately or other aspects connected to experimental error.

Section C

Question 10

(a) There were some very good answers to this question. The majority of answers were Level 2 and gave good details of sample types of question and question content. Answers achieving Level 3 marks included excellent details particularly of sample and sampling methods, but also an awareness of the specific ethical issues that would apply to this study. Weaker scripts suggested experiments, observations or other methods that did not directly respond to the question. In such instances, there were often unnecessarily described aspects of their study such as experimental design, IVs and DVs, using time which would have been better spent writing about details of the questionnaire. For example, many gave muddled descriptions of question formats, e.g. saying ‘open question’ but then giving an example of a closed question without options rather than a question that elicits detail. Similarly with examples of closed questions; many lacked answer options. There was also quite a lot of evaluation included in some answers.

(b) There were some good answers to this question and very few rubric errors. Some answers were quite generic but there were some very detailed ones suggesting appropriate ways of overcoming limitations.

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 able to recognise terminology, theories, disorders and studies identified in the syllabus as well as key details of named theories and studies as some were unable to identify and/or define the terms/theories given in these type of questions. Creating a glossary of key terms, revision of terminology/theories using flash cards and class quizzes on terminology/theories could prove useful. These questions are worth 2 marks and a brief response is appropriate.

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

These questions could ask the candidate to describe a theory, study or treatment(s) used by psychologists that are 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 could be given the title used in the syllabus, for example, Schizophrenia and psychotic disorders: treatment and management: cognitive-behavioural therapy (Sensky, 2008).

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

These questions could require the candidate to explain up to two strengths and/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. 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, theory or technique(s) 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 could 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)

This question will always come from one of the bullet points in the syllabus. Candidates could describe the three (or four) studies, theories, explanations of disorders 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, explanations disorders 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)

This question will always ask the candidate to evaluate the theories, studies, explanations of disorders 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 tended to achieve higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the theories, studies, disorders and techniques described in the **part (a)** of the answer. Responses 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 this question is worth 10 marks and attempt to include an appropriate amount of information.

General comments

The marks achieved by candidates for this session of the 9990 specification achieved across the full range of the mark band which was very pleasing to see. Many scripts showed good knowledge, understanding and evaluation throughout their responses. Some limited knowledge and understanding with brief and/or superficial responses. These responses often had limited evaluation skills.

Time management for this paper was good for the majority and most scripts answered all questions that were required. A few 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) There were many good responses to this question with an outline given of how Freeman used virtual reality to assess the symptoms of schizophrenia. Most responses were able to identify that the virtual reality was set in a library/train/underground. Some responses outlined how the symptoms were assessed such as the comments made by the participants about the avatars or identified one or more of the measures used by Freeman such as Green paranoid thought scale. Weaker responses identified that the virtual reality was in a library or train but did not give any details of how the symptoms were assessed or gave vague details such as behaviour was noted down which was not creditworthy.

(b) There were some clear and somewhat detailed responses describing the study by Sensky on cognitive-behavioural therapy (CBT) as a treatment for schizophrenia. Common details of the study given in responses included the sample used, the two treatment groups (CBT and be-friending) as well as some responses giving details of these therapies, details of how and when symptoms were assessed and a result from the study. Weaker responses gave fewer details of the study. Some responses gave details on the sample and procedure of the study but then did not give any of the results. Very weak responses outlined CBT as a treatment but gave no details of the study. A few responses outlined the Lovell et al. study on using CBT face-to-face versus CBT over the telephone as a treatment for obsessive-compulsive disorder which was not creditworthy.

(c) The responses to this question covered the full range of the mark scheme. Better responses identified two strengths and explained these with clear examples from the Sensky study. Common strengths included the use of random allocation to conditions, use of a control group, longitudinal method which allowed long-term effects to be seen, strengths of CBT itself. Other common strengths included that as quantitative data was used comparisons can be made, effectiveness of CBT (over be-friending) and ethics. Responses that used these strengths tended to achieve fewer marks due to lack of clear examples. Some responses gave a strength that the nurses who did the therapies were well-trained. This was not a strength of the study as any practitioner who does therapy will have been trained. There were also a number of responses that identified issues such as reliability and validity but gave no explanation of how or why these issues could be applied to the Sensky study and be considered to be a strength and therefore received no credit.

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. Some responses were detailed, accurate and coherent with a good use of psychological terminology. The strongest responses focused on describing the three explanations of obsessive-compulsive disorder (OCD) including biomedical, cognitive and behavioural and psychodynamic. Most responses gave very detailed descriptions of the genetic, biochemical and neurological with many able to identify genes, hormones and areas of the brain involved in the development of OCD. There were some very strong responses that were able to explain the relationship between oxytocin dysfunction and symptoms of OCD as well as why abnormalities in the basal ganglia could lead to OCD. There were also some strong descriptions of the cognitive-behavioural explanation with a good outline of the faulty cognitions/obsessions in OCD and how these can lead to compulsions. Some responses outlined how the compulsions can act as negative and positive reinforcement which led to repetition of the compulsive behaviour. Others gave a general outline of anxiety disorder, rather than OCD specifically, and how anxiety has a cognitive-behavioural explanation.

Weaker responses often gave a description of either the cognitive explanation or the behavioural explanation. Some of these did not link the cognitions or the reinforcement to the symptoms of OCD. There was a variety of details given for the psychodynamic explanation. A few gave a lot of detail of the anal stage of development and the conflict that can arise at this stage as well as how it could be unresolved and therefore create symptoms of OCD in later life. Weaker responses tended to identify the anal stage of development and that this could cause someone to be anally retentive/ expulsive but were then not able to explain why being retentive/expulsive could lead to OCD. Some responses outlined the symptoms of OCD and/or case studies of individuals with OCD (e.g. Charles) but this was not answering the question and was therefore not creditworthy.

(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 explanations 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 explanation rather than applying the issue to the different explanations. A few responses did effectively discuss the named issue of reductionism versus holism, but this rarely had any analysis. Most responses attempted the named issue but there were many that were confused about the meaning of reductionism and holism or simply stated that one explanation ignored the other explanations which provided very limited evaluation. With the named issue, the majority of responses outlined reductionism as failing to consider other explanations rather than it being about breaking behaviour into simpler units.

Other common evaluation issues included nature versus nurture, individual and situational, scientific nature or not of the explanation and determinism. Higher mark responses gave examples to back up their points whereas the weaker responses often identified how the explanation fit with the evaluation issue (e.g. it is on the nature side of the debate) with no details of why or any analysis to suggest that the explanation could also be considered from the other side of the debate/approach.

Psychology and Consumer Behaviour

Question 3

(a) There were many strong responses to this question with a clear definition of overload and good examples of how shoppers might experience overload. Some responses explained that overload of personal space in a shop could cause the shopper to leave whereas others explained that the shopper would become aware of the sight, smell, etc. of other customers due to their close proximity. A few responses believed that overload occurred due to the shop having too many items or displays within it. This would not lead to an overload of personal space as this involves having our personal space bubble invaded by other people and not by displays in a shop. A significant number of responses did not give an example from shoppers.

(b) There were some clear and somewhat detailed responses describing the study by Robson et al. about table spacing in restaurants. Most achieved full marks usually by identifying the sample, the 3 conditions of the 2 IVs and some findings. Somewhat weaker responses gave fewer details often with some incorrect information about the independent variables or omitted any findings from their description. Some responses stated the study was carried out in a real restaurant which was incorrect and therefore not creditworthy.

(c) The vast majority of responses were able to identify and sometimes describe with examples one strength and one weakness of the study by Robson et al. Common strengths included generalisability as sample is large and usefulness to restaurants. Common weaknesses included lack of ecological validity, problems with self-reports and cultural bias. The majority of responses were not developed and achieved limited credit. Some responses stated that the ecological validity of the study was good and the control in the study was poor or there were problems due to observer bias. None of these points were creditworthy.

Question 4

(a) Many responses achieved Level 2 or 3 for this question. Good descriptions were often seen in responses for the Fischer et al. study as well as the Auty and Lewis study on product placement in films. Many responses gave a summary of the conclusions of the Snyder and DeBono study with limited details of any of the three studies they conducted. The descriptions of Kohli et al. on effective advertising tended to be brief and sometimes vague with only one or two of the recommendations given by Kohli described in the response.

(b) A few excellent responses were given that explained how Fischer et al. investigated using a technique that children would be happy to carry out (a game), reducing any potential distress. These responses were also able to provide a good discussion regarding the ethics of studying children and how the Fischer et al. study got consent from the parents which improved the ethics of the study. Weaker responses focussed on ethical problems with the Fischer et al. study but often could not explain why the study did not follow the guidelines and/or was not required to follow the guideline. These responses simply stated the study did not get consent from the children and distressed the participants. Other common issues were generalisability and usefulness. Weaker evaluation often identified the sample used in the study and then stated it could not be generalised with no explanation given as to why not. In addition, the response would state that a study was useful and then re-state the results of the study without giving any specific applications for the study. Stronger responses explained why children, American participants, etc. might produce different results to older children or people from non-Western cultures.

Psychology and Health

Question 5

(a) This was often well answered with many responses achieving full marks with good outline of what is meant by ‘chronic pain’. Many responses stated that it is long lasting pain that can last 3 months or more. Responses often gave examples of chronic pain such as migraines, arthritis and cancer pain. A very small minority of responses outlined acute pain rather than chronic pain. These type of responses were not creditworthy.

(b) The vast majority of responses identified acupuncture and TENS. Some responses extended this by outlining how each alternative technique leads to a reduction in pain. Some responses explained that acupuncture restores the flow of Qi or releases endorphins which reduce pain. Good responses for TENS outlined that TENS can relax muscles and/or release endorphins leading to pain reduction. Weaker responses gave vague, confused or incorrect details of how these alternative techniques reduce pain. Responses that were not creditworthy outlined psychological techniques such as imagery that can reduce pain.

(c) The responses to this question covered the full range of the mark scheme. Better responses outlined both a similarity and a difference between a medical technique and an alternative technique for managing and controlling pain. Differences were better answered than similarities with the most common, strong response being that medicine has side effects and the alternative does not. Responses often identified the side effects of the medication but usually just stated the alternative does not have side effects without explaining why this is true. Creditworthy similarities included that a practitioner needs to provide the treatment and the cost of treatment. Weaker responses often stated that both treatments reduce pain without any explanation about how they reduce pain.

Question 6

(a) The responses to this question covered the full range of the mark scheme. Stronger responses had a good understanding of the research into practitioner and patient interpersonal skills (non-verbal communications and verbal communications). The studies by McKinstry and Wang, together with McKinlay, were frequently described extremely well. Weaker responses gave fewer details of the research or some of the details were incorrect. These type of responses often achieved in Level 1 or Level 2. A significant minority of responses wrote good accounts of studies but did not include any results. Some studies were overly detailed, candidates should be encouraged to write appropriately for the mark allocation, for instance following the model of 'sample, design, procedure, results'.

(b) There were some good responses to this question. The names issue of generalisability was handled well with a number of responses citing how the sample in McKinstry and Wang was large, increasing generalisability, but they were all from the same region of Scotland, reducing generalisability. Cultural bias was also used as well as issues such as validity (including temporal validity) and practical applications. There were very few good examples of sound analysis in responses. Some responses outlined a number of issues but in a superficial manner, simply stating that issues did or did not apply to studies but not why. These type of responses often achieved in the Level 1 mark band.

Psychology and Organisations

Question 7

(a) There were many full mark responses to this question. The majority were able to name two roles (either specific – shaper and plant were common choices) or the general roles such as action-oriented or people-oriented. The command word 'Identify' means to simply name yet a high proportion of responses described the roles they had identified. Just writing 'shaper, plant' was sufficient for full marks. A common mistake made in some responses was to identify types of followers such as sheep and yes man which was not creditworthy.

(b) Most responses were able to achieve full marks by identifying the correct five stages of the theory and giving a brief outline of each stage. Weaker responses identified fewer stages and often put these stages into the incorrect order and/or gave incorrect descriptions of the stages. The most common error was to confuse 'storming' and 'norming'. A small minority of responses did not know Tuckman's theory and gave an anecdotal description of groups within an organisation which was not creditworthy.

(c) The responses to this question covered the full range of the mark scheme. Better responses outlined both a strength and a weakness of Tuckman's theory with an example to back up each point. The most common weaknesses given were about a lack of timescale and that group development can be cyclical rather than linear. A common strength was how useful the theory is with more sophisticated responses explaining how managers will be able to understand that 'storming' is a normal part of group development and conflicts do not mean that the group cannot

work together successfully. Weaker responses often stated as the weakness that the theory would not apply to all groups but failed to explain why. A common strength was that the theory is useful to organisation but no details were given of how an organisation could use this information. These type of responses usually achieved a Level 1 mark.

Question 8

(a) There were some good, detailed responses to this question with clear detailed research that investigated health and safety in organisational work conditions. There were some detailed descriptions of the Three Mile Island accident. Fox et al. and Cowpe were most commonly described well, with Fox et al. being more comprehensive on the whole. Some very good descriptions. There were also a few descriptions of Gold et al., which were creditworthy. Most of the details given of real-life accidents at work tended to be vague or incorrect information given compared to actual studies. A significant minority of responses did not give any results in their account of studies. Weaker responses tended to give very brief details of the studies, often just stating that token economies and/or adverts reduced accidents and chip pan fires with no details of the studies given.

(b) Most responses to this question evaluated the studies using the named issue of experiments. Many answers simply identified that the studies were field studies and provided either very limited evaluation of field experiments, or just explained why the study could be considered a field experiment. Some responses did identify and sometimes outline (with examples) the strengths and weaknesses of field experiments. Popular points included good ecological validity, poor control and reliability. In addition to experiments, popular evaluation issues were generalisability and practical applications. Ethics was a popular evaluation issue raised in responses but this issue was usually covered in a very superficial way by stating whether the study met a guideline with no indication/ example of how the study met or broke the guidelines. Some responses suggested that the Fox et al. study was harmful to the participants but could not explain why they thought this about the study. In addition, some responses attempted to raise evaluation issues which were not appropriate to this topic such as nature versus nurture which was not creditworthy.

PSYCHOLOGY

Paper 9990/42
Paper 4 Specialist Options: Application

Key messages

(a) 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.

(b) Questions should be read carefully ensuring that the focus is on what the question asks.

(c) All components of the question should be included in answers. For example, **Question part (d)** for **Questions 1**, **2**, **3** and **4** required advantages and disadvantages (plurals), examples of each and a conclusion.

(d) 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.

(e) 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. Description cannot be credited.

(f) Psychological knowledge should be applied wherever possible. Anecdotal and common-sense answers will not achieve top marks.

General comments

Not all scripts selected two options, with a few answering questions from one option only. Some answered questions from three and even four options. Whilst answers to one option were often very good, some answers to the second option were poor, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

(i) Candidates are advised to read the 'stem' of the question, the introduction or the opening words in Section A questions as the information provided is crucial to answering each question part that follows.

(ii) Answers must refer to the study the question is about. Many answers provided general comments which were unrelated to the study itself.

(iii) For question **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so restricted marks.

(iv) Many conclusions merely 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.

(v) 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.

(vi) Candidates should always provide sufficient detail to score all the available marks. A single sentence is more likely to score 1 mark rather than 2 marks, so a little elaboration, for example, that goes beyond the basic sentence is always recommended. Candidates should always try to display psychological knowledge where possible.

Section B

Many responses attempted to conduct an experiment whatever the question, even when an interview, questionnaire or observation, or other methods independent of an experiment were required. Candidates should not try to make other methods 'fit' into an experimental format. 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 a topic area is insufficient to score full marks. 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. Answers that don't 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) Most responses were awarded full marks for explaining that MAOIs (or monoamine oxidase inhibitors) inhibit oxidase and that this results in serotonin and dopamine being maintained at high levels and so reducing depression. Some confused MAOIs with SSRIs or did not know the function of MAOIs.

(b) (i) Many responses did not display knowledge of what a randomised control trial was, or they confused it with random sampling.

(ii) One strength of an RCT is that researchers do not allocate participants to groups, it is done randomly, and so there is no bias and this potential extraneous variable is removed. Answers stating this were awarded full marks.

(c) All candidates were awarded at least 1 mark for suggesting a treatment for depression other than MAOIs. Some opted for SSRIs but answers were often very brief. Others chose electroconvulsive therapy and answers were much more detailed, often scoring full marks. Other candidates opted for a psychological therapy, with the work by Beck and Ellis featuring.

(d) Most responses had few problems with providing a range of strengths and weaknesses in relation to the use of drugs. However, these were often not related to depression.

Question 2

(a) Any relevant participant variable was credited. However, there needed to be some elaboration for full marks and responses stated nothing more than 'age and sex'. Full mark answers were typically 'some participants may have preferred salty foods'; some participants may not have liked some of the foods being tasted'; 'some participants may have been smokers, which affects taste'.

(b) (i) To be awarded 2 marks, answers needed to include two components. First there was a need to write more than '*questionnaire*', '*scale*' or '*rating scale*'. Strong answers mentioned a '*Likert-type*' scale ranging from +7 to –7. Second, what was being measured in Experiment 1 was sweetness, saltiness and liking, and not crunchiness (which was Experiment 2).

(ii) A number of responses incorrectly referred to crunchiness. An answer scoring full marks would be 'reported sweetness and saltiness was significantly lower in the loud compared to the quiet sound conditions' and this was written by many.

(c) (i) Some responses did not appear to know the term counterbalancing and nearly always provided incorrect responses. Some explained the term but did not relate this to the study, as the question required, so did not score both marks.

(ii) Like **part (c)(i)** many answers included an effect if the conditions were not counterbalanced, namely order or practice effects, but did not always relate this to the study by Woods et al. If a question states 'in this study' then the answer must be shown to be related to the study.

(d) Many answers included two strengths and two weaknesses of controls, but often only scored partial marks because frequently answers were not related to the study by Woods et al. as the question required. An appropriate answer would be 'controls reduce extraneous variables so the level of noise had to be the same for all participants. In the Woods et al. study the level of loud noise was always between 75 and 85 decibels.' Some weaker responses began focusing on controls, but ended with a focus on experiments, sometimes making points which were not related to controls.

Question 3

(a) The reason why it was important to measure unhealthy food consumption in addition to healthy food consumption was first to have a baseline of what children were eating and second, to allow a comparison to be made over time (e.g. at the end of the study) to know if the programme had achieved a change in eating behaviour. Many scored full marks for their answers.

(b) There was some misunderstanding about the term 'food consumption techniques' despite this being the term used by Tapper et al. Most responses achieved marks by writing about relevant aspects. Three techniques were used in the Tapper et al. study: taste exposure, modelling and rewards. This meant, for example, that a candidate explaining how children achieved rewards, such as the use of stickers, etc. were awarded 2 marks.

(c) Many responses did not address all parts of the question. Weaker responses wrote about informed consent in detail, but made no mention of the Tapper et al. study (as the question required). The same was true of other ethical guidelines such as confidentiality, right to withdraw and debriefing.

(d) This question focused on health promotion strategies in schools, and so any answer which included alternative strategies (such as fear arousal) or used examples that were not based in schools scored no marks. The range of strengths and weaknesses was often very good, but like **part (d)** answers for other options an appropriate conclusion was not always provided.

Question 4

(a) Many responses achieved full marks. Weaker answers were vague or did not display knowledge of the term.

(b) (i) Any need that was justified as 'individual' could be credited. However, whilst physiological needs (each *individual* needs food, warmth) were found easy to justify, social needs (experience of belongingness) proved to be much more difficult. Answers which merely identified a need scored no marks, and neither did answers which listed all five needs.

(ii) Any need that was justified as 'situational' could be credited. Like question **part (b)(i)** some needs were found easier to justify than others and social needs being situational was most common here.

(c) Many responses achieved full marks. Explaining that Maslow has 5 (or 8) needs reduced to Alderfer's 3 was credited. Any explanation of how this reduction was achieved was also credited. For example, Maslow's physiological and safety needs were reduced to existence needs by Alderfer and similarly reductions by Alderfer resulted in relatedness needs and growth needs.

(d) Answers to this question part followed the same pattern as **part (d)** answers in other options. In this case advantages and disadvantages were good, but were frequently not related to any need theory of motivation. Any advantage or disadvantage (or strength or weakness) *must* be related to the question being answered in order to achieve marks.

Section B

Question 5

(a) There were many excellent answers, but also some answers that were confused and incorrect. For example, a few confused imaginal desensitisation with covert sensitisation and had designs where participants were inappropriately vomiting rather than merely leaving the scene. Some answers chose participants with gambling or pyromania as their ICD, but then assessed them (their DV) using the K-SAS (the kleptomania symptom assessment scale) which is inappropriate for gambling or pyromania. Additionally, a number of responses suggested using YBOCS to assess the effectiveness of the treatment. YBOCS measures obsessive-compulsive disorder not impulse control disorders (although PG-YBOCS could be used if the disorder being studied was gambling).

(b) The psychological evidence often saw imaginal desensitisation being confused with covert sensitisation and this resulted in no marks being awarded. Similarly use of K-SAS scored no marks if the participants did not have kleptomania. Some candidates described the study by Glover and this was inappropriate because it was about the use of covert sensitisation. Responses about Blaszczynski and Nower often scored some marks, but often did not explain how that study related to the design of their study. Finally, it should be noted that having a large sample does not necessarily make a sample more generalisable (a small 10 or large 1000 students are still students). Making the sample more representative might make it more generalisable.

Question 6

(a) As is often stated, an experiment is a specific method and includes features specific to it. Other methods such as questionnaires, interviews and observations are not experiments and have their own specific features. Responses should not try to make experimental features fit to non-experiment methods. For this question the required method was a questionnaire, yet many experiments were designed at the expense of essential features of questionnaires. Some correctly devised questionnaires but often focused on pre-cognitive decisions rather than intentions and behaviour.

(b) Relevant psychological evidence was that of Ajzen (1991) and the theory of planned behaviour. This is where a person might say they intend to buy a product (and often does so) but sometimes they do not buy the product. Methodologically a questionnaire was an appropriate method so a shopper's thoughts could be accessed at the intentions stage and then again post-purchase stage and if the product was not purchased the reasons for the change in thought process could be investigated. A number of responses instead focussed on the work of Knutson et al. appearing to interpret the question as relating to pre-cognitive decisions.

Question 7

(a) There were many strong answers in response to this question with the specific features of experiments being evident. To improve, responses should make sure that the IV is appropriate to answer the question and that the DV is an appropriate measure of the IV. Most candidates compared biofeedback with an alternative stress reduction technique or had a 'no feedback' control condition. A small number of responses confused the IV and DV, and a small number of candidates did not display knowledge of what biofeedback was, restricting the marks awarded.

(b) The study by Budzynski et al. featured as psychological knowledge and many responses described the study accurately. However, marks were often restricted because answers did not explain how the Budzynski et al. study had informed their design. Methodologically answers were good, and candidates should ensure that explanations for their design decisions are provided rather than simple statements. For example, '*I used a self-selecting sample*' is not an explanation.

Question 8

(a) See opening comments for **Question 6(a)** which also apply here. Many responses were not able to display knowledge of the term adaptive leadership. Moving office, for example, is not *organisational* change and an adaptive leader would not be needed to facilitate that change. For those responses using a questionnaire, questions about job satisfaction were often asked, which is peripheral to adaptive leadership. The most apposite questions to ask workers would be about 'getting on the balcony', 'regulating distress' or any other feature of adaptive leadership. Doing this would show how psychological knowledge informed the design.

(b) Following on from **(a)** above, psychological knowledge should have been about adaptive leadership and how that knowledge was used to design the questionnaire given to workers. Instead, the features of adaptive leadership were described and not related, and often any theory of leadership was included, such as Fielder or autocratic and democratic styles.

Section C

Question 9

The strongest responses were able to consider whether behavioural explanations of impulse control disorders were too reductionist to be useful. Answers scoring marks at the bottom end of the mark range often *described* nothing more than a number of explanations of ICDs, with little reference to reductionism. A number of responses focused on 'useful' at the expense of the rest of the question and their answers focussed on whether an explanation is useful or not. Answers like this struggled to provide any argument, usually nothing more than 'this explanation is useful because it applies in the real world' as a treatment for ICDs.

Question 10

There were a few very good answers in response to this question which were based on appropriate psychological knowledge, mainly that of North et al. Such answers presented arguments both for and against generalisations and used the North et al. study in support of the arguments. At the bottom end of the mark range, answers made general points such as the inability to generalise from one culture to another and such answers often included no reference to any psychological knowledge at all.

Question 11

There were a few very good responses that applied the use of questionnaires to unrealistic optimism and showed clear use of Weinstein and his study on unrealistic optimism. Most answers were weaker, although the debate on questionnaires was good, because other than using the words '*unrealistic optimism*' responses did not show knowledge of what the term was referring to.

Question 12

The focus of this question should have been on the statement given, with a debate both for and against situational leadership. Whilst a few responses did this, many others focused on describing theories of leadership, without focus on the statement, and sometimes scored no marks. The focus of questions in this section is evaluation rather than description, with credit achieved for agreeing or disagreeing with the statement and using evidence in support.