PSYCHOLOGY

> **Paper 9990/11**
> **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. 'In depth' tends to mean having two examples of a particular concept or to support an evaluative point. 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 explain similarities and/or differences between studies based on psychology. Brief, rudimentary responses can rarely be credited.

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.

There was a number of blank responses (every single question had blank responses). As positive marking is used, candidates should attempt all questions even if they are unsure of the answer they are providing.

Cambridge Assessment
International Education

© 2023

Comments on specific questions

Question 1

(a) A minority of responses correctly identified the group. It is important for candidates to clearly label which answer they are choosing on multiple choice questions, especially if they change their initial response. There were multiple instances where a candidate had chosen more than one answer.

(b) Stronger responses could clearly outline the category of 'ignores' from the named study. Common points made by candidates included the participant making no verbal comments and only working on their questionnaire. However, the majority of responses were not correct. Common errors included writing about the 'ignorant group' in the study or giving a generic definition of 'ignore'. When a question states 'in this study' the response must be focused on the named study.

Question 2

(a) Features of a sample are any element related to the sample used in the named study. There were many correct responses to this question with popular choices being the number of participants, the gender imbalance, and the participant panel. Common errors included identifying features from a different study and providing incorrect features of the sample used in the study by Andrade.

(b) The question required an identification of a strength, so no explanation was needed. Common correct responses focused on the potential generalisability of the sample or that they were already fatigued/bored. There were a minority of responses that provided a correct strength of the study but not the sample so could not be awarded credit. It is important for candidates to read questions carefully.

Question 3

(a) The majority of responses could describe part of the Disgust/Fear hierarchy. Popular choices included the that it was a list of different buttons and the scale used to rate each of these buttons. There were some responses that clearly described three points about the hierarchy. However, some responses presented information about the therapy sessions or attempted to describe the difference between disgust and fear, neither of which could be credited. Again, it is crucial for candidates to read questions carefully.

(b) There were many correct responses to this question in terms of at least identifying and explaining a weakness of the hierarchy. Popular choices included subjectivity and social desirability. However, there were a significant minority of responses that explained a correct weakness but did not explain why it was a weakness using evidence from the study. It is important for candidates to note the number of marks assigned, and the number of problems that need to be addressed (e.g. one), as this typically represents how the marks can be awarded. In this type of question, one mark is for identifying a potential weakness and the second mark must come from explicit example(s) from the study.

Question 4

(a) Stronger responses could clearly describe four aspects of the victims used in the study. Common choices included the role of the 'drunk' victim and the 'ill' victim and the clothing worn by them. However, there were several errors presented by candidates. There was a significant minority that described the 'ill' victim as being an old man and the 'drunk' victim as always young, plus some candidates claimed there were female victims. Other errors included describing the sample of participants or what the victim had to do during a trial. Neither were answering the question set.

(b) There were some clear, detailed conclusions provided by candidates to this question, including how helping is affected by type of victim and gender differences in helping behaviour. There were instances where candidates wrote the same conclusion as was in the question so no credit could be awarded. Also, there were some candidates who provided a result. Candidates need to know the difference between a result and a conclusion (see key messages).

Question 5

This question had the most blank responses on the paper. The majority of candidates could not provide results based on the two questionnaires named in the question. Stronger responses however could present a full result with data comparing either between-groups or within-group for either or both questionnaires. Common errors included describing what was measured by each questionnaire or presenting results that were not known (e.g., number of participants choosing a rating).

Question 6

The minority of responses could outline two assumptions of the cognitive approach. Popular choices focused on the retention of information and the way we process information (input-process-output). Stronger responses could then use an explicit example from one of the core studies from the cognitive approach to show how it supports that assumption. Some candidates did provide an example from a study but it was not linked to the presented assumption so no credit could be awarded. There were multiple examples of candidates providing assumptions from a different approach, mainly biological. Assumptions are on the syllabus and need to be known by candidates. This question had the second highest blank response rate on the paper.

Question 7

The minority of responses could suggest two real-world applications based on Milgram. Popular choices included teachers dressing authoritatively or using an authority figure like a doctor to present messages about health campaigns. Like with all questions related to real-world applications, it is important for candidates to outline what the application is and then explain how it can be achieved using evidence from the named study, in this case Milgram. Additionally, the 'what' part needs to be explicit, therefore responses that stated 'to become more obedient' could not be credited as this is not explicit. Some responses described the study by Milgram with no application or suggested future research that could be based on the study. Neither of these could be awarded credit. Finally, there were a number of responses that explained certain behaviours in real-life rather than a real-world application. These could also not be credited. Responses need to be prospective and not retrospective. There were very few responses that were not ethical.

Question 8

Stronger responses could clearly provide one reason why each of the people in the question is correct in terms of validity in relation to Baron-Cohen et al. Popular examples for Danilo included high levels of internal validity and that the revised version was designed to have increased validity. Stronger responses could provide specific and explicit examples from the study to support the reason. Popular examples for Noah included low levels of ecological validity and that people could still 'guess correctly'. Again, stronger responses could then provide specific and explicit examples. Weaker responses tended to provide more than one reason and in this case the strongest 'one reason' was credited. To improve, candidates need to have examples from each core study that appropriately support core concepts on the syllabus like validity.

Question 9

(a) The majority of responses could describe at least two features of REM sleep. Popular choices included it being Rapid Eye Movement and that dreams occur more often in this phase of sleep. Incorrect responses included describing the content of dreams reported in the study by Dement and Kleitman or claiming we only dream in this phase of sleep.

(b) Stronger responses could clearly explain one similarity and one difference. Popular choices to compare the studies on included generalisability of the sample, how brain activity was measured, and the collection of quantitative data. To improve responses to this type of question, candidates need to choose comparison points that can be developed and explained, using examples from both studies to explain the similarity and/or difference. For example, explaining the different samples can focus on the characteristics and/or sampling technique and allows for a detailed response with a statement about generalisability. However, stating that each study had a different aim does not allow the response to be detailed. Candidates need to choose their comparisons carefully to ensure that they are logical and can be explained fully.

Cambridge Assessment
International Education

Question 10

The strongest responses evaluated the study by Bandura 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 reliability. Common choices included generalisability, reliability, ecological validity, ethics, and quantitative data. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Bandura 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 Bandura 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. 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 responses appeared to follow a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). Some responses appeared to be prepared essays for Bandura et al. without one of their points being about reliability. A response that does not 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 reliability 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. In this series, more responses were attempting to focus on real-world application which tended to only be awarded partial credit as this question is evaluative in nature and not application. To be awarded credit for 'application', the candidate must present an evaluative strength and then, as a result of this strength, explain what positive real-world application can be seen as a consequence of the strength presented (e.g. good for classification of films/computer games because Bandura et al. found…). To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, avoiding real-world application where possible. 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. An evaluation that is in depth tends to have at least two explicit examples from the named study for every evaluative point made. These are the requirements for a Level 4 response.

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. 'In depth' tends to mean having two examples of a particular concept or to support an evaluative point. 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 explain similarities and/or differences between studies based on psychology. Brief, rudimentary responses can rarely be credited.

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.

There was a number of blank responses (every single question had blank responses). As positive marking is used, candidates should attempt all questions even if they are unsure of the answer they are providing.

Comments on specific questions

Question 1

(a) A minority of responses correctly identified the group. It is important for candidates to clearly label which answer they are choosing on multiple choice questions, especially if they change their initial response. There were multiple instances where a candidate had chosen more than one answer.

(b) Stronger responses could clearly outline the category of 'watches' from the named study. Common points made by candidates included the participant making no verbal response to the stooge and only observing the stooge. However, the majority of responses were not correct. Common errors included giving a generic definition of 'watches'. When a question states 'in this study' the response must be focused on the named study. This question had the highest rate of blank responses on the paper.

Question 2

(a) The majority of candidates could name at least one test used in the study by Baron-Cohen et al. The most common choice was the Autism Spectrum Quotient. A common error was to identify the gender in the picture, which is not a test, just a recognition exercise. A minority of candidates identified the eyes test, but this was in the question so could not be credited. Candidates need to read questions carefully.

(b) The question required an identification of a strength, so no explanation was needed. Common correct responses focused on the comparability of quantitative data or that the test was standardised and reliable. There were a minority of responses that provided a correct strength of the study but not the test so could not be awarded credit. Again, it is important for candidates to read questions carefully.

Question 3

(a) The majority of responses could describe part of the procedure for measuring emotional arousal. Popular choices included the rating scale and that the participants had to press a button for their response. There were some responses that clearly described three points about the measurement. However, some responses presented information about the scanning procedure or the recognition task, neither of which could be credited. Again, it is crucial for candidates to read questions carefully.

(b) There were many correct responses to this question in terms of at least identifying and explaining a weakness of the measurement of emotional arousal. Popular choices included subjectivity and it used a forced-choice technique. However, there were a significant minority of responses that explained a correct weakness but did not explain why it was a weakness using evidence from the study. It is important for candidates to note the number of marks assigned, and the number of problems that need to be addressed (e.g. one), as this typically represents how the marks can be awarded. In this type of question, one mark is for identifying a potential weakness and the second mark must come from explicit example(s) from the study.

Question 4

(a) Stronger responses could clearly describe four aspects in relation to the allocation of the children to the experimental groups. Common choices included the pre-rating of aggression by teachers and that a matched pairs/triplets approach was used. However, there were several errors presented by candidates. There was a significant minority that described how many participants per group in great detail, but the question was about how they were allocated not what the groups were as this was in the question.

(b) There were many clear, detailed conclusions provided by candidates to this question, including how gender affected aggression imitated and that behaviour can be imitated later in the absence of the model. There were instances where candidates wrote the same conclusion as was in the question so no credit could be awarded. Also, there were some candidates who provided a result, rather than a conclusion. Candidates need to know the difference between a result and a conclusion (see key messages).

Cambridge Assessment
International Education

© 2023

Question 5

The majority of candidates could not provide results based on the two questionnaires named in the question. Stronger responses, however, could present a full result with data comparing either between-groups or within-group for either or both questionnaires. Common errors included describing what was measured by each questionnaire or presenting results that were not known (e.g. number of participants choosing a rating), or incorrect comparisons. For the Food Preferences Questionnaire there were many responses comparing the Love Group and the Control Group. This never happened as it was an analysis of Believers and Non-believers. It is important for candidates to know the main results from all measures used in any core study on the syllabus. This question had the second highest blank response rate on the paper.

Question 6

The minority of responses could outline two assumptions of the social approach. Popular choices focused on the role of either individuals or groups on our behaviours, emotions, and cognitions. Stronger responses could then use an explicit example from one of the core studies from the social approach to show how it supports that assumption. Some candidates did provide an example from a study but it was not linked to the presented assumption so no credit could be awarded. There were multiple examples of candidates providing assumptions from a different approach, mainly learning (social learning was a common error). Assumptions are on the syllabus and need to be known by candidates.

Question 7

The minority of responses could suggest two real-world applications based on Dement and Kleitman. Popular choices included helping to diagnose sleeping disorders and promoting better quality sleep. Like with all questions related to real-world applications, it is important for candidates to outline what the application is and then explain how it can be achieved using evidence from the named study, in this case Dement and Kleitman. Additionally, the 'how' part needs to be explicit, therefore responses that stated 'using an EEG to see if a patient has a sleeping disorder' could not be credited as this is not explicit. Some responses described the study by Dement and Kleitman with no application or suggested future research that could be based on the study. Neither of these could be awarded credit. Some candidates suggested two real-world applications but simply named two different sleep disorders but could only be credited for the best one (this is an example of repetition; the two suggestions must be distinctly different). Finally, there were a number of responses that <u>explained</u> certain behaviours in real-life rather than a real-world <u>application</u>. These could also not be credited. Responses need to be prospective and not retrospective. There were very few responses that were not ethical.

Question 8

Stronger responses could clearly provide one reason why each person in the question is correct in terms of ethics in relation to the study by Yamamoto et al. Popular examples for Haji included using the smallest number of chimpanzees and that they were socially housed. Stronger responses could provide specific and explicit examples from the study to support the reason. Popular examples for Dottie included the use of rewards for just one chimpanzee and that choosing the incorrect tool might cause distress. Again, stronger responses could then provide specific and explicit examples. Weaker responses tended to provide more than one reason and in this case the best 'one reason' was credited. A minority of candidates attempted to argue about the chimpanzees not being in the wild but this could not gain credit as the chimpanzees were laboratory-raised. Also, there were a minority of candidates who argued it was not ethical based on guidelines for human participants, for example, informed consent, which clearly could not gain credit. It is important for candidates to know the ethical guidelines in relation to animals and apply them to studies like Yamamoto correctly. To improve, candidates need to have examples from each core study that appropriately support core concepts on the syllabus like ethics.

Question 9

(a)   The majority of responses could outline what a phobia is. Popular choices included mentioning irrational fear and that it is towards an object/stimulus/situation. Incorrect responses included writing about the role of disgust, naming different types of phobias, or stating it is fear of 'something' which is too generic and not based on psychology.

(b)   The majority of candidates could not provide a correct outline of evaluative learning. Many responses appeared to focus on it being an <u>evaluation</u> of a process (looking for strengths or

Cambridge Assessment
International Education

weaknesses) or that it is operant conditioning. Stronger responses could clearly outline the role of classical conditioning and cognitive elements in this type of learning.

(c) Stronger responses could clearly explain one similarity and one difference. Popular choices to compare the studies on included the lack of generalisability of the sample, the use of the case study method, and the collection of quantitative data. To improve responses to this type of question, candidates need to choose comparison points that can be developed and explained, using examples from both studies to explain the similarity and/or difference. For example, explaining the similarity of the case study method using examples from both studies and then explaining why it is a similarity could be awarded Level 4. However, stating that each study had a different aim or used a different species does not allow the response to be detailed. Candidates need to choose their comparisons carefully to ensure that they are logical and can be explained fully.

Question 10

The strongest responses evaluated the study by Andrade 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 generalisability, reliability, ecological validity, ethics, and quantitative data. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Andrade 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 Andrade 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. 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 responses appeared to follow a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). Therefore, some responses appeared to be prepared essays for Andrade without one of their points being about independent measures. A response that does not 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 reliability 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 or could only provide a strength or weakness in general without any examples from the study by Andrade. In this series, more responses were attempting to focus on real-world application which tended to only be awarded partial credit as this question is evaluative in nature and not application. To be awarded credit for 'application', the candidate must present an evaluative strength and then, as a result of this strength, explain what positive real-world application can be seen as a consequence of the strength presented (e.g. good for jobs that are repetitive but require focused attention as Andrade found…). To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, avoiding real-world application where possible. 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. An evaluation that is in depth tends to have at least two explicit examples from the named study for every evaluative point made. These are the requirements for a Level 4 response.

PSYCHOLOGY

> **Paper 9990/13**
> **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. 'In depth' tends to mean having two examples of a particular concept or to support an evaluative point. 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 explain similarities and/or differences between studies based on psychology. Brief, rudimentary responses can rarely be credited.

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.

There was a number of blank responses (every single question had blank responses). As positive marking is used, candidates should attempt all questions even if they are unsure of the answer they are providing.

Comments on specific questions

Question 1

(a) A minority of responses correctly identified the group. It is important for candidates to clearly label which answer they are choosing on multiple choice questions, especially if they change their initial response. There were multiple instances where a candidate had chosen more than one answer.

(b) Stronger responses could clearly outline the category of neutral from the named study. Common points made by candidates included an irrelevant response given to the stooge. However, the majority of responses were not correct. Common errors included giving a generic definition of 'neutral'. When a question states 'in this study' the response must be focused on the named study. This question had the second highest rate of blank responses on the paper.

Question 2

(a) The majority of candidates could name at least one of the measures observed in the study by Piliavin et al. The most common choice was the race. A common error was to identify features of the victim. Candidates need to read questions carefully.

(b) The question required an identification of a strength, so no explanation was needed. Common correct responses focused on ecological validity or that the sample was generalisable. There were a minority of responses that provided a correct weakness of the study, but the question was asking about a strength. Again, it is important for candidates to read questions carefully.

Question 3

(a) The majority of responses could describe a small part of the procedure for the final part of the study by Bandura et al. Popular choices included what was in the room and the response categories that were observed. There were some responses that clearly described three points about the measurement. However, some responses presented information about the initial part of the study, or the part where the children got intentionally frustrated. Again, it is crucial for candidates to read questions carefully.

(b) There were some correct responses to this question in terms of at least identifying and explaining a weakness of the measurement of the final part of the study by Bandura et al. Popular choices included ethics and validity of recording. However, there were a significant minority of responses that explained a correct weakness but did not explain why it was a weakness using evidence from the study. It is important for candidates to note the number of marks assigned, and the number of problems that need to be addressed (e.g. one), as this typically represents how the marks can be awarded. In this type of question, one mark is for identifying a potential weakness and the second mark must come from explicit example(s) from that part of the study. There were many responses that gave a generic weakness of the overall study rather than the final part of the study.

Question 4

(a) Stronger responses could clearly describe four aspects in relation to the feedback from the learner. Common choices included that the responses were predetermined and that nothing was heard until Shock Level 300v. However, there were several errors presented by candidates. Many responses gave generic ideas about the procedure without any specific detail (e.g. the shock levels at which certain activities happened) and so could only be awarded minimal credit.

(b) There were many clear, detailed conclusions provided by candidates to this question, including obedience being linked to diffusion of responsibility and that authority figures will be followed. There were instances where candidates wrote the same conclusion as was in the question so no credit could be awarded. Also, there were some candidates who provided a result, rather than a conclusion. Candidates need to know the difference between a result and a conclusion (see key messages).

Question 5

The majority of candidates could not provide results based on the two questionnaires named in the question. Stronger responses, however, could present a full result with data comparing either between-groups or within-group for either or both questionnaires. Common errors included describing what was measured by each questionnaire or presenting results that were not known (e.g. number of participants choosing a rating), or incorrect comparisons. For the Photograph Ratings there were many responses comparing the Love Group and the Control Group. This never happened as it was an analysis of Believers and Non-believers. It is important for candidates to know the main results from all measures used in any core study on the syllabus. This question had the highest blank response rate on the paper.

Question 6

The minority of responses could outline two assumptions of the learning approach. Popular choices focused on the role of social learning and operant conditioning in behavioural changes. Stronger responses could then use an explicit example from one of the core studies from the learning approach to show how it supports that assumption. Some candidates did provide an example from a study but it was not linked to the presented assumption so no credit could be awarded. There were multiple examples of candidates providing assumptions from a different approach, mainly social. Assumptions are on the syllabus and need to be known by candidates.

Question 7

The minority of responses could suggest two real-world applications based on Baron-Cohen et al. Popular choices included helping to diagnose AS/HFA or that the test can help improve social intelligence. Like with all questions related to real-world applications, it is important for candidates to outline what the application is and then explain how it can be achieved using evidence from the named study, in this case Baron-Cohen et al. Additionally, the 'how' part needs to be explicit, therefore responses that stated 'give children the eyes test at school' could not be credited as this is not explicit. Some responses described the study by Baron-Cohen et al. with no application or suggested future research that could be based on the study. Neither of these could be awarded credit. Finally, there were a number of responses that explained certain behaviours in real-life rather than a real-world application. These could also not be credited. Responses need to be prospective and not retrospective. There were no responses that were not ethical.

Question 8

Stronger responses could clearly provide one reason why each person in the question is correct in terms of ethics in relation to the study by Saavedra and Silverman. Popular examples for Maria included obtaining informed consent from the mother and boy, plus there was some level of confidentiality. Stronger responses could provide specific and explicit examples from the study to support the reason. Popular examples for Oscar included causing psychological distress by being exposed to buttons, plus some privacy may have been breached. Again, stronger responses could then provide specific and explicit examples. Weaker responses tended to provide more than one reason and in this case the best 'one reason' was credited. It is important for candidates to know the ethical guidelines in relation to humans and apply them to studies like Saavedra and Silverman correctly. To improve, candidates need to have examples from each core study that appropriately support core concepts on the syllabus like ethics.

Question 9

(a)   The majority of responses could identify one element of altruism. Popular choices included doing something that may disadvantage yourself. Incorrect responses included giving examples of altruism or confusing altruism with empathy.

(b)   The majority of candidates could not provide a correct outline of targeting helping. Many responses appeared to be tautological in nature, stating that targeted helping was helping in a specific/targeted way. Stronger responses could clearly outline the role of appropriate help given based on understanding the situation.

(c)   Stronger responses could clearly explain one similarity and one difference. Popular choices to compare the studies on included the collection of quantitative data and sampling. To improve responses to this type of question, candidates need to choose comparison points that can be developed and explained, using examples from both studies to explain the similarity and/or difference. For example, explaining the similarity of collecting mainly quantitative data using

examples from both studies and then explaining why it is a similarity could be awarded Level 4. However, stating that each study had a different aim or used a different species does not allow the response to be detailed. Candidates need to choose their comparisons carefully to ensure that they are logical and can be explained fully.

Question 10

The strongest responses evaluated the study by Dement and Kleitman in depth and in terms of two strengths and two weaknesses, with at least one of these points covering the named issue of reliability. Common choices included generalisability, reliability, ecological validity, ethics, and quantitative data. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Dement and Kleitman 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 Dement and Kleitman 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. 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 responses appeared to follow a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). Some responses appeared to be prepared essays for Dement and Kleitman without one of their points being about reliability. A response that does not 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 reliability in an evaluative sense but did not fully explain <u>why</u> it could be a strength and/or a weakness. Several responses did not cover the named issue or could only provide a strength or weakness in general without any examples from the study by Dement and Kleitman. In this series, more responses were attempting to focus on real-world application which tended to only be awarded partial credit as this question is evaluative in nature and <u>not</u> application. To be awarded credit for 'application' the candidate must present an evaluative strength and then, as a result of this strength, explain what positive real-world application can be seen as a consequence of the strength presented (e.g. good for helping to diagnose sleep disorders because Dement and Kleitman found…). To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, avoiding real-world application where possible. Each strength and weakness should be of equal length with an explanation as to <u>why</u> it is a strength or weakness with examples from the study to show clear understanding. An evaluation that is in depth tends to have at least two explicit examples from the named study for every evaluative point made. These are the requirements for a Level 4 response.

PSYCHOLOGY

| Paper 9990/21 |
| Paper 2 Research Methods |

Key messages

This question 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. These types of questions require candidates to use a variety of skills. Candidates should be encouraged to prepare for each of these skills, especially for demonstrating knowledge of concepts and the application of this knowledge.

Ability to apply knowledge and understanding to novel scenarios is essential to help candidates to successfully complete this paper. This skill can help candidates in two ways: candidates should 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 should be aware of questions which require a link. When a question includes 'in this study', or makes a direct reference to the scenario, responses should go beyond simply describing or evaluating. The answer must also be contextualised in a relevant way. Practice could help candidates to learn both how to extract relevant ideas, and how to make novel suggestions based on scenarios.

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

Question 10 in this paper requires candidates to produce an original design for a novel research question. This 'creative' process requires practice and it is important 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**) candidates should have had the experience of practical problems in conducting studies. This is a high-level skill and can be developed through practical work with designing and conducting small studies in class, or through practice with novel scenarios. Candidates should be familiar with the overall structure of **Question 10(a)**, which can be closely tailored to requirements of an individual question, such as the required research methods and the scenario.

Comments on specific questions

Section A

Question 1

(a)   Most candidates were able to explain the term 'ecological validity' which is 'the extent to which the findings of a study set in one situation will generalise to other situations'. An answer using these words or similar would be awarded 1 mark. The second mark was available for any example which illustrated ecological validity. The most common example was the Piliavin et al. 'Subway Samaritans' study, being high in ecological validity, but also credit was awarded to the candidates giving an example of a study which was low in ecological validity, such as that by Milgram.

Question 2

(a)   Many candidates were awarded only 1 mark because their answer was limited to 'both closed'. The question asks for an explanation and candidates stating a reason why the items were both closed, such as the details in the stem of the question, scored full marks.

(b) This question consisted of two parts: firstly, an explanation of what is meant by the term 'critical item' and secondly an example from the Laney et al. study to show understanding of the term 'critical item'. Nearly all candidates knew that asparagus was the critical item, and were awarded 1 mark, but not all candidates could explain why this was 'critical'. A critical item is something that is directly related to the aim, something that is related to the IV or is a crucial part of the procedure.

Question 3

(a) A number of candidates described an independent measures design, which is incorrect. Many others correctly described a repeated measures design which is where every participant does all the conditions of the independent variable. The second part of this question required an example from the Canli et al. study. Only a few candidates gave the correct answer of the independent variable being the emotive/negative and non-emotive/neutral stimuli presented in the first part of the study. The recognition task, done three weeks later, was not part of the independent variable. A common error: a repeated measures design is not repeating the same condition again (which is replication) rather, it is participating in all the conditions of the IV.

(b) Those candidates describing an independent measures design in **part (a)** also scored no marks for this question part. However, most candidates gave an appropriate strength of a repeated measures design to achieve 1 mark and for the second available mark many elaborated on the strength or provided an appropriate example from a core study.

Question 4

This question required two 'ways' for four marks, so candidates only providing one way could only score a maximum of two marks. Examples of correct ways included: the number of items was the same (1 mark) and included stick, straw, belt, etc. (2 marks); the time period to give the straw or stick was always the same (1 mark) 5 minutes before the researcher intervened (2 marks). Other possibilities related to the room, the familiarisation before starting the trials or any other standardised aspect of the procedure could be credited.

Question 5

Many candidates appeared to confuse reliability with replication i.e. referring to repeating a study and getting the same results rather than the consistency of any test or task or measure. In this instance, about reliability in experiments, several answers were acceptable. For example, if tests are inconsistent, differences between conditions might be due to poor measurement, not the independent variable. Also acceptable was that if the DV is not being measured reliably then the data cannot test a causal effect. A few candidates referred to inter-rater reliability which is most frequently used in observation studies, rather than in experiments.

Question 6

There were many correct answers which received full marks. Amongst the incorrect answers, many claimed that 'receiving a reward' or 'monetary payment' was a feature of volunteer sampling. However, many answers also claimed this for opportunity sampling. Receiving money or course credits can apply to both and is not a specific feature of one technique or the other. Many answers stated that 'a feature of volunteer sampling is that participants volunteer'. Such answers received no credit for repeating the term without elaboration. Some answers stated that volunteer sampling involves placing an advertisement in a newspaper. This correct answer would get 2 marks.

Question 7

(a)(i) Despite the question stating the requirement of two situational variables, many candidates wrote about two 'personal' variables. Hence, answers such as 'the participant's mood' and 'their upbringing', were included, all of which were awarded 0 marks. A number of situational variables also scored 0 marks, because they could apply to all types of shop. For example, 'the weather', external to the shop, might affect a person's mood, but this is before they enter the shop and so not creditworthy. Many answers with situational variables did receive credit, such as the number of people in the shop (crowding), and the temperature inside the shop.

(ii) In **(a)(ii)**, credit was given for an explanation of how the situational feature might affect frustration. For example, crowding inside the shop may frustrate a customer who cannot see or get to the goods they want to purchase. Answers that did not refer to frustration received no credit.

(b)(i) Most candidates were awarded 1 mark for stating that right to withdraw could not be given because participants did not know they were in a study. The second mark could only be awarded if the answer referred to the 'in this study' component of the question, and most answers did not do this.

(ii) Similar **to (b)(i)**, question **(b)(ii)** required a link to the study in the question and here the answer needed to make a reference to privacy, such as 'shoppers might be purchasing something very personal so observing this without consent would break the guideline of privacy'. Answers simply restating the same as for **part (i)** - that privacy could not be achieved because participants did not know they were in a study - received no credit.

Question 8

(a) This question required candidates to suggest an objective measure. At a simple level, this meant constructing a measure to produce numerical data. However, the question stated that a self-report rating scale could not be used, so candidates suggesting a rating scale scored 0 marks. Candidates also scored 0 marks for suggesting an open-ended questionnaire or suggesting asking a question that resulted in a yes or no answer (this does not produce quantitative data). Success at learning a language could be measured with a test looking at the number of correct responses, or the number of errors. It could also be measured by the size of vocabulary (number of words learned). To achieve the second available mark candidates had to justify their answer. Most candidates appeared to struggle to do this.

(b) Very few candidates scored the available mark due to omitting one aspect of the following, to calculate a median for enjoyment of word puzzles in this study the scores/ratings need to be placed in numerical order and the middle score needs to be identified.

(c) Most candidates achieved the maximum four marks for their answers to this question. Marks were awarded for labelling each axis ('language learning score' and 'enjoyment of word puzzles'), for labelling the word puzzle axis from 0–5, for plotting data showing a scatter graph, for the data showing a positive slope, and for including a line of best fit (which must be a straight line). A few candidates drew a bar chart which scored limited credit for the labelling of the axes.

Question 9

(a) In response to this question, many candidates scored 0 marks for stating 'a directional hypothesis shows the direction'. A statement like this shows no understanding of what a directional hypothesis is. A correct answer for limited credit might be 'because Inma knows which way the results will go'. For the second available mark, the answer needed to relate to Inma's experiment which involved stating the IV and the DV. A full-mark correct answer would be 'because she believes that the animals will be faster (to learn the maze) in the light than the dark'.

(b) There are only two ways to write a null hypothesis, and applied to Inma's experiment these are: 'there will be no difference between speed in the maze in the light and in the dark' or 'any difference between maze learning success in the light and in the dark is due to chance'. Either of these null hypotheses scored the one available mark. Any other wording is incorrect.

(c) This question required an advantage of using an independent measure design compared to using a repeated measures design. This meant that any answer had to include both designs in order to achieve full marks. Correct answers included 'unlike a repeated measures design there will be no fatigue/practice effects (1 mark) that are created by the animal going through the maze again' (2 marks). This answer has the required contrast and it is related to Inma's study.

(d) Ethical guidelines for animals are different from ethical guidelines for humans and so any answer that included informed consent, deception or right to withdraw was awarded 0 marks. Correct ethical guidelines could include housing, reward, no deprivation and no pain or distress. The answer also had to be related to Inma's experiment and included in the stem of the question was the statement 'If the animal does not find the food after 10 minutes, they are given food by the experimenter' which led most candidates to link this to the guideline of reward, no deprivation or no harm.

Question 10

(a)   This question required candidates to design an observational study. This meant that inclusion of a number of essential features of an observation were required, such as whether the observation was structured or unstructured, participant or non-participant, naturalistic or controlled, and covert or overt. In addition, to achieve a high mark candidates could consider behavioural categories and how these could be operationalised, whether time sampling or event sampling would be applied and whether there would be 2 observers so the consistency of the observations could be checked (inter-rater reliability). All of these needed to be applied to the behaviour of candidates in a library at a university. Also relevant was who the participants would be and the location of the study on the sampling technique. Answers achieving full marks included all of these features, clearly and in detail. Answers including some of these features achieved mid band marks and answers including very few or none of these features were awarded very low or no marks at all. Some candidates include the use of questionnaires which were irrelevant to the question. Some candidates tried to apply the features of an experiment, but this was also irrelevant. If the method stated is an observation, that method is required, and an alternative method cannot be credited. The features of an experiment should only be applied if the question specifically asks candidates to design an experiment.

(b)   A number of candidates suggested that ethics was a limitation or that the sample was limited, but ethics and sampling were excluded in the question. Any limitation related to experiments also scored 0 marks. Candidates awarded full marks often referred to a limitation of just one observer not being able to observe every behaviour (for example) and a solution to this being to have two observers so the consistency of the observed data could be checked. Notably, many candidates stated that having two observers improves reliability, but it merely checks whether there is consistency or not. Another possible limitation was that the candidates in the library might notice they were being observed by Dr Clare and the suggested solution was to use CCTV cameras and analyse the data at a later date.

Cambridge Assessment
International Education

© 2023

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 skills.
- Candidates demonstrated excellent knowledge of basic concepts such as identifying independent and dependent variables and basic strengths of sampling methods. The analytical questions, and expanding on answers, they found more challenging. Some areas of basic concepts where knowledge was less evident was in candidates' ability to identify and suggest ways to measure variables. Here, candidates' responses often lacked accurate description and/or examples.
- The ability to link accurate detail to a given scenario or context is also required on the paper. This more difficult skill was tackled well by some candidates, but less well by many others.

General comments

Candidates were able to access marks across the whole paper. However, not all were able to accurately and/or consistently demonstrate knowledge and understanding or to access the additional marks for linking their response to the scenarios, thus limiting their performance as a whole. 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 some straightforward questions such as **Questions 1**, **3(a)(i)**, and **9(a)**, **9(b)** and **9(c)** than on more demanding ones, such as **Questions 7(c)(ii)**, **8(a)(i)**, and **10**. However, there were also some more straightforward questions which candidates found challenging, such as **Questions 5**, **7(a)(i)** and **(ii)**, **7(b)(ii)** and **8(b)**. Where questions required a link, for example to a study, candidates were often able to earn partial marks for an initial identification of a relevant fact, such as identifying the disadvantage in **Question 7(b)(ii)** but were then unable to relate this explicitly to the information in the stem. **Question 10** was sometimes well answered, although responses often lacked any indication of the specific details of an interview, such as the type of interview, type of questions and examples of questions. Such responses therefore lacked necessary relevant detail.

As a general point, candidates can always make use of blank paper within the booklet or extra pages but it is advisable to indicate this at the end of a response that is continued. All pages are checked but on occasions it is difficult to establish to which question extra material belongs unless candidates have made this clear.

Comments on specific questions

Section A

Question 1

In general, this question was answered well, with candidates demonstrating a good understanding of the purpose of the familiarisation stage of the Yamamoto study.

Question 2

(a)   There were a range of answers provided for this question, with stronger responses demonstrating an excellent understanding of a non-directional hypothesis although weaker responses often predicted that one type of conditioning would be faster than the other or referred to relationships/ correlations rather than differences/effects. Some linked to phobias and other behaviours.

(b) Generally not well answered, with many candidates simply repeating 'the same' from the question and not offering any further explanation of this. Many answers suggested that Chloe should have used a one tailed rather than a two tailed hypothesis, which was not addressing the question. Stronger responses were less common here, but explained that as there was no difference, the hypothesis would need to be rejected.

Question 3

(a) (i) Candidates have a good understanding of the deceptions in this study, most commonly the deception of telling participants they were getting a Suproxin injection when really it was adrenaline / a placebo. Some candidates only offered part of the answer, for example stating that participants were told they were being given Suproxin, without explaining why this was a deception / what was really happening.

(ii) Most candidates were able to explain that deception reduced demand characteristics. However, not all candidates provided their answer in the context of this specific study and these responses scored 1 mark. For the second mark, candidates needed to explicitly link the concept of demand characteristics to this study.

(b) Most candidates referred to lowered validity / inability to record data but some misunderstood and suggested why the problem had occurred and not why this was a problem.

Question 4

(a) Most candidates were able to draw a bar graph, although labelling was sometimes incorrect and some candidates drew histograms rather than bar charts and so were not able to access full marks.

(b) This was a challenging question for many, with the majority of the candidates providing a description of the data rather than a conclusion. It is important that candidates distinguish between results (for example, that the mean rating was higher in week 2 than week 1) and conclusions (for example, that the critical item changed the beliefs of participants, or that it is possible to implant false memories of liking asparagus).

Question 5

This proved to be a challenging question for many candidates. Many candidates gave a great deal of detail from the Bandura study without directly addressing the question, and there are still candidates misunderstanding the concept of inter-rater reliability as simply having more than one observer, rather than the level of agreement between multiple observers. Stronger candidates were able to explain that having good inter-rater reliability meant that all observers had the same understanding of aggression and this means that the differences between the groups are much more likely to be due to differences in the children's behaviour rather than differences in the way the behaviour was observed.

Question 6

Strong answers showed a good understanding of the need to control extraneous variables and the differences between laboratory and field experiments in terms of their ability to do this. Weaker answers focused on control groups and control (manipulation) of independent variables. Strong answers gave detailed examples from studies such as Canli et al. and Andrade for laboratory experiments and Piliavin et al. for field experiments.

Section B

Question 7

(a) (i) Although the majority of candidates correctly identified the IV as gender, there were a significant number who thought that the IV was the number looking in the window, or even some characteristic of the window such as its size or how clear it was.

(ii) As with **7(a)(i)**, the majority were correct in identifying the DV, but gender and other suggestions were also seen with some frequency.

(b) (i) There were some very strong responses to this question, many containing substantially more detail than was required for 3 marks. Strong answers explained that a set of categories would be predetermined and then gave examples of these. However, some candidates appeared to misunderstand the question and explained other aspects of the observation, such as watching from behind a screen or the need to gain consent.

(ii) Although the majority of candidates understand that the disadvantage of a structured observation is being unable to record behaviours that do not fit the categories, most gave this generic answer without linking this explicitly to this study. A small number of strong responses gave an example of a behaviour that was not included in the categories and explained that this would not be recorded.

(c) (i) Most candidates were able to identify that the median is the 'middle' number when numbers are listed from lowest to highest, although many repeated 'central' from the question and were not awarded any marks.

(ii) Most candidates offered an appropriate alternative measure although the second mark was less frequently awarded. Either responses were simply generic and offered no link to the study at all, or were not creditworthy for other reasons, such as suggesting that the mode could be used to see whether males or females looked in the window more often.

Question 8

(a) (i) This question produced a wide variety of responses. Some candidates did not offer any appropriate suggestions for measuring vocabulary size, and simply suggested mean / median etc. Some misunderstood the question and offered suggestions relating to the physical size in terms of centimetres. Some candidates also seemed to struggle to understand the term 'vocabulary' correctly, which impacted their marks.

However, there were lots of appropriate suggestions made with some very creative ideas, ranging from asking for definitions, to open ended questions on a variety of topics or asking candidates to write stories. Stronger answers included explicit details of how these would produce numerical data, such as the number of correct answers out of ten, the number of synonyms produced in a minute etc, whereas weaker responses simply offered the idea with no indication of how this would produce a score.

(ii) Many answers simply said that the data would be quantitative, which was not enough for a mark as quantitative data would need to be produced for a correlation. Stronger answers said that tests were standardised or that marking would be objective.

(b) There was also some misunderstanding of what was required here. It was not enough to simply say that participants would be asked for their level of education, the answer needed to show how the level of education would be measured. Other misunderstandings were to suggest that tests of vocabulary or IQ could be used to assess level of education. The most common responses were to suggest number of years of education completed or current grade.

(c) (i) The majority of candidates were able to identify this as a negative correlation although some did identify this as positive.

(ii) This graph showed no correlation and although the majority of candidates were able to identify this, a significant number used terms such as 'difference' or 'effect' in their answers here.

Question 9

(a) Generally well answered, with most candidates saying that the advantage was that it was easy/quick because participants are available.

(b) Generally well answered with most candidates understanding that random sampling produces a more diverse, and therefore more representative sample. Some candidates did however, give descriptions of random sampling which did not address the question.

(c) Very well answered with most candidates identifying increased generalisability.

Section C

Question 10

(a)   There were some very strong responses to this question where candidates gave detailed information about the type of interview, type of questions and examples of questions. However, there were also some who simply described the procedure used by Dement and Kleitman without recognising the need for the response to focus on the interview procedure. Where candidates did this, they focused on aspects of laboratory experiments such as the IVs and DVs, controls such as not having caffeine or alcohol and the equipment (EEG and EOG) at the expense of giving details about an interview. This meant that there was little information provided on the three 'majors' required for this question. There seems to be some misunderstanding of 'semi-structured' interviews with some candidates appearing to think this means a mix of open and closed questions, rather than a mix of prepared questions and questions asked in response to participant responses. However, most candidates offered appropriate discussion of ethical issues.

(b)   The marks candidates achieved here depended on how they had responded to **part (a)**. If they had focused on an experimental procedure then weaknesses to this was not creditworthy here, with the answer needing to focus on the interview aspect of the procedure. However, when candidates did focus explicitly on the interview, a range of strong suggestions for improvements were made. As always, despite the question explicitly telling candidates not to refer to sampling or ethics, some answers did exactly that.

Cambridge Assessment
International Education

PSYCHOLOGY

> **Paper 9990/23**
> **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 skills.
- Candidates demonstrated excellent knowledge of basic concepts such as generalisation and naming experimental designs. The analytical questions, and expanding on answers, they found more challenging. Some areas of basic concepts where knowledge was less evident was in candidates' understanding of random sampling, measures of spread, and inter-rater reliability. Here, candidates' responses often lacked accurate description and examples.
- The ability to link accurate detail to a given scenario or context is also required on the paper. This more difficult skill was tackled well by some candidates, but less well by many others.

General comments

Candidates were able to access marks across the whole paper. However, the issues identified above limited performance for some individuals on some questions.

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 straightforward questions such as **2** (ethical guidelines for animals), **4(b)** (categories in Pepperberg), **5** (differences between REM and nREM sleep), **7(a)(i)** and **(ii)** (conditions in experiments), **8(a)** (techniques for data collection), and **9(a)(i)** and **(ii)** (questions on a questionnaire). This was especially when they were low mark tariff, such as questions **7(a)(i)** and **9(a)**. However, where candidates were asked to link their answer to a context, they were less able to apply their knowledge successfully.

Question 10 was sometimes very well answered although many responses were muddled. Either the candidate designed a study that was not a correlation, or their measurement of one or both variables was incomplete or inappropriate. These are both essential details for the description of a study.

There were very few papers with several questions for which the candidate provided no response at all.

Comments on specific questions

Section A

Question 1

(a) The concept of random sampling was poorly understood and/or poorly expressed. Responses were typically circular, saying the people were selected 'randomly' / 'at random' or that they 'didn't know who they were'. Other incorrect responses were descriptions of opportunity sampling. Stronger answers referred to selecting by chance, or all members of the population having an equal possibility of being selected. There were also some competent descriptions of how this could be done.

(b) The answers to this question part were stronger than for part **(a)**, with some candidates able to explain, for example, why representativeness could not be guaranteed.

Question 2

Responses to this question were fairly successful, but many candidates were unable to identify – by name or description – guidelines relating to the use of animals. In such cases, there was either no reference to guidelines or those that were described were not guidelines relating to animals. Those candidates who offered the best responses tended to choose the numbers and housing guidelines.

Question 3

(a) There were very few correct answers here. Where the mark was not earned, the candidate had typically given a circular response, using the term 'spread'.

(b) Some candidates earned the mark for stating one other measure of spread (the range), though fewer were able to clearly describe an advantage of standard deviation. A small number of candidates explained an advantage of using the standard deviation in isolation from an alternative, and were able to earn this mark alone.

Question 4

(a) Many students were able to make an attempt at suggesting that it was because these categories have 'lots of choices' in them but were less often able to expand on this suggestion to earn credit for detail.

(b) Most candidates gave a creditworthy answer here, typically saying 'size', although there was a range of alternatives offered.

Question 5

Although this question was well answered, there was still a significant proportion of candidates who did not earn full marks. The most common correct points were more dreaming in REM than nREM, more REM at the end but more nREM at the start, rapid eye movements in REM but not in nREM. Many candidates incorrectly stated that 'REM sleep is deeper than nREM'.

Question 6

This question was moderately well answered, with some candidates achieving good marks. Most candidates were able to describe the importance of making participants aware of the nature of the study so that they could choose to participate, though fewer candidates went on to explain how this might be achieved, (e.g. through briefing of participants by asking for written agreement). Examples of gaining consent often did not contain sufficient detail to earn marks, although examples where consent was not obtained were more likely to be creditworthy.

Descriptions of protection from harm were tackled more successfully, with more detail and examples. References to inappropriate guidelines were rare.

Section B

Question 7

(a) (i) This question part was well answered. Many candidates gave one of the two obvious responses, but there was a range of other creditworthy suggestions. The most common mistake was to make suggestions of conditions that would only have been suitable for a different study, i.e. one testing a different aim. These were not appropriate responses as conditions for Nigel's study.

(ii) This question part was also well answered, with most candidates gaining at least one of the two possible marks. Candidates who did not earn credit here typically seemed to struggle with the concept of a control or of a comparison.

(b) Some candidates were able to give good explanations, although fewer gave successful examples. Where examples were given, they were usually correct but sometimes the candidate had not attempted this part of the question.

(c)   This question part was well answered, with many candidates gaining the mark. Where candidates could name an alternative design they gained the mark.

Question 8

(a)   This question part was generally well answered, although in some responses there was insufficient detail for one, or both, suggestions to earn full credit. For example, a minority of candidates just named techniques, or offered very brief descriptions such as 'by asking questions'. In other cases, the suggestions were generic, rather than being related to the scenario given.

(b)(i) Most candidates earned some credit for their response to this question, but many were unable to explain why their suggestion was advantageous, i.e. they were unable to elaborate on their answer. For example, an advantage of observations could be that you might gather more information by watching than the participant is able to express. This is correct, and creditworthy, but needs detail to earn full credit, for example by suggesting that the participant may not want to admit that they are confused, but the observer might see them frowning, or the same idea could be given in a generic way, simply by stating that the observer could gain extra information from facial expressions.

(ii)   As above, although the candidates appeared to find this question part somewhat easier.

(c)   This question part was well answered.

Question 9

(a)(i) This question part was often very well answered. However, many candidates only earned one mark as they did not indicate the possible 'answers' participants could give, which is essential for a closed question. For example, alternative answers should be given (such as yes/no) and for rating scales the scoring must indicate the meaning of high and low values. A minority of candidates suggested irrelevant questions.

(ii)   This question part was also well answered, although here candidates sometimes offered questions that would not elicit descriptive answers, effectively giving closed questions without the possible answers. A true open question invites elaboration, such as using 'describe' or 'explain'. A minority of candidates also suggested irrelevant questions here.

(b)(i) This question part was generally well answered, with most candidates scoring at least one mark.

(ii)   As with part **9(b)(i)**, there were some good attempts for this question part, with candidates scoring at least one mark. However, candidates were less successful in this question part than in **9(b)(i)**.

(c)   Responses to this question were very variable, with candidates scoring the range of possible marks. In other words, some candidates were unable to answer the question at all, whilst others gave accurate and detailed answers. This question was also the most commonly omitted by candidates.

Section C

Question 10

(a)   This question part appeared to prove challenging for lots of candidates. Many candidates designed experimental studies rather than correlational ones and some collected data that could not be correlated. Study procedures were typically muddled, either in terms of the participants being used or how the variable of animal size was being measured. Severity of the phobia was typically measured, and described, in a more effective way. Many candidates needed to make more specific mention of ethics in direct relation to exposing people with phobias to their feared objects.

10(b)   This question part also appeared to prove challenging. As many candidates designed experimental rather than correlation studies, responses here were often irrelevant. Many candidates did not respond to the instruction to suggest how the study could be done differently.

PSYCHOLOGY

Key messages

Key messages

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

These questions are worth 2 marks and it is important that candidates produce a brief response. The focus here should be on terminology, theories, and disorders identified in the syllabus, to inform knowledge and understanding of these short answer questions. Some of the responses were very strong, achieving full marks, however, some were unable to identify and/or define the terms/concepts given in this type of question. Revision aids and continued testing could support candidates to maximise their mark.

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

These questions are worth 4 marks and require a more detailed outline of a theory, study or treatment(s) used by psychologists 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 and a result, or a summary of the key features of the study. Candidates should be encouraged be able to identify which part of the syllabus the question is referring to.

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

These questions are worth 6 marks and require the candidate to explain either two strengths, two weaknesses, or one strength and one weakness. These questions are linked to the answer that they have described in **part (b)**. The responses should be in some detail showing a clear understanding of the question. For full marks there will be a strength and a weakness, two strengths, or two weaknesses, providing a good explanation with clear detail. Candidates need to be able to apply what they understand about research methods to the studies outlined in the syllabus. Many excellent responses were seen which contained detailed strengths/weaknesses with clear examples from the study listed in the question. Other responses needed to develop more detailed and better-balanced arguments to achieve higher marks. For example, some responses were generic and not specific to the study, theory or technique(s) named in the question. Stronger responses gave specific examples in detail to support their points with some balance.

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

These questions are worth 8 marks and 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 syllabus under the appropriate bullet point. It is important that candidates are clear of the studies from the syllabus, so as not to use the incorrect topic area. Candidates need to outline at least two of the studies, theories, explanations of disorders or techniques in some detail to achieve full marks. Stronger responses demonstrated accurate and detailed descriptions with excellent understanding of the question. Other responses needed to develop more detailed and better-balanced arguments to achieve higher marks.

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

These questions are worth 10 marks and ask the candidate to evaluate the theories, studies, explanations of disorders and/or techniques described in **part (a)** of the question. There is a named issue that should be included. The named issue could be better addressed at the start of the answer to meet the requirements of the question. Alongside this, there should be at least two evaluation issues in some detail to have a range of issues for full marks. It would be advisable to include a conclusion at the end of each issue to develop the

response further. 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. Many excellent responses contained detailed evaluation points whereas other responses needed to develop more detailed and better-balanced arguments to achieve higher marks and analysis needs to be more sophisticated. 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. 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 counterargument to the issue under discussion or comparing the issue between two studies and/or theories. Evaluation must be comprehensive, and the range of issues covered should be highly relevant to the question for top band answers.

General comments

Candidates were able to make good use of the time allocated for this series and most candidates attempted all questions. Almost all candidates answered the appropriate topics and where this was not the case, the candidate attempted to answer all the topic areas with limited success.

The most popular topic was Psychology and Abnormality followed by Psychology and Health.

Candidates appeared to be well prepared for the examination series, with particularly high levels of knowledge and understanding of the topic areas. Some candidates appeared to be less prepared and appeared to struggle with the questions, especially in terms of evaluation skills. These candidates often gave generic, superficial responses which were brief.

Overall, candidates for this session were well prepared, showing good knowledge and understanding and evaluation throughout. For some candidates more preparation is required, focusing on detail and evaluation and analysis skills.

Comments on specific questions

Psychology and Abnormality

Question 1

(a)   Some good knowledge and understanding of the term pyromania. Most responses stated that pyromania is a desire to set fire to things and gained 1 mark for this brief point. Stronger responses were able to gain a second mark for the point that the person feels anxiety / heightened arousal prior to setting the fire and once they have done it the arousal reduces, or related it to a fascination with fire.

Very few candidates wrote nothing creditworthy, except for those who confused pyromania with another disorder.

(b)   This question required candidates to describe the study by Glover (2011). Strong responses gave a detailed answer with clear understanding of the topic area. Details such as sample, procedure, number of sessions and results of the study by Glover (2011) that used covert sensitisation as a treatment for kleptomania were stronger responses. Other responses understood fewer of the details of the study and or confused some of the details.

(c)   The responses to this question covered the full range of the mark scheme. Strong responses identified a strength and a weakness of the study by Glover (2011) in some detail to highlight the points made. Most popular responses for the strength were in-depth data / qualitative data and effectiveness of covert sensitisation as a treatment for kleptomania. For the weakness the most popular response was poor generalisability. Ethics for the weakness was not considered to be creditworthy here, since the treatment was required, and consent given. To maximise marks, candidates should be encouraged to state the strength or weakness in some detail and link to the study by Glover to highlight why this makes a strength or a weakness.

Question 2

(a) This question required candidates to describe explanations of depression, using genetic and neurochemical (Oruc et al, 1997), cognitive (Beck, 1979) and learned helplessness / attributional style (Seligman, 1988). The stronger responses covered all three in some detail. Weaker responses often gave limited descriptions with significant errors. The study by Seligman was often confused with the 1965 study with dogs rather than the learned helplessness / attributional style study. For the genetic and neurochemical explanation, some candidates discussed neurochemical explanations (low levels of serotonin). This was credited but better responses outlined the Oruc study in some detail. 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.

(b) This question required candidates to evaluate explanations of depression, with the named issue being about reductionism versus holism. There were many Level 1 and 2 responses due to limited analysis. Candidates overall did engage with the named issue. Other responses focused on the nature/nurture debate or the deterministic nature of the studies. The strongest responses covered this question by structuring it by issues, along with supporting examples from the studies in **2(a)** and analysis. Candidates should be aware that if the named issue is not addressed, a maximum of 5 marks can be awarded and if only the named issue is addressed, a maximum of 4 marks can be awarded.

Psychology and Consumer Behaviour

Question 3

(a) This question was answered well, with responses correctly identifying the spatial behaviour movements / patterns of movement of the study by Gil et al. (2009). Some candidates used a study other than Gil et al., which was not creditworthy as it did not answer the question set.

(b) This question required candidates to outline any two results from the studies by Dayan and Bar-Hillel. Strong responses gave a detailed answer with clear understanding of the results of the study, using descriptions or percentages. Weaker responses appeared to struggle to recall any result, particularly study 2.

(c) This question required candidates to discuss the validity of the studies by Dayan and Bar-Hillel. The most common responses included ecological validity and population validity. Top band answers gave clear understanding of the question and were able to explain at least two points regarding validity with links to either study 1 or 2. Weaker responses tended to state the point about validity but needed to include elaboration from the study by Dayan and Bar-Hillel.

Question 4

(a) This question required candidates to describe what psychologists have discovered about retail/leisure environment design using Turley and Milliman (2000), Finlay et al. (2006) and Vrechopoulos (2004). The stronger responses covered all three in some detail. Weaker responses often gave limited descriptions with significant errors. 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. There were a lot of Level 2 responses due to the description lacking detail and some adequate use of psychological terminology. Highly detailed descriptions of two or more of the studies can achieve top band.

(b) This question required candidates to evaluate what psychologists have discovered about retail/leisure environment design, with the named issue being generalisability. There were many Level 1 and 2 responses due to limited analysis. Candidates overall did engage with the named issue whereas other responses focused on the practical applications / usefulness. The strongest responses covered this question by structuring it by issues, along with supporting examples from the studies in **4(a)** and analysis.

Psychology and Health

Question 5

(a) Candidates often appeared to struggle to state the type I and type II errors the correct way around. There were many 1-mark answers for stating 'false negative' or 'false positive' without an example or giving a basic outline of type I and type II errors without an example. Stronger responses were when candidates give a description of a type I and type II error with an example.

(b) Many candidates gave detailed descriptions of the procedure of the study by Robinson and West (1992), including the sample/location. Many candidates cited other similar studies which were not given credit. For questions such as these where the procedure is required, results were not creditworthy.

(c) The most likely weaknesses given were lack of generalisability and social desirability bias. Top band answers required candidates to show a clear understanding of the question and explain two weaknesses with clear links to the study by Robinson and West (1992). Many responses would give a weakness but not in any detail or without any link to the study. For example, just stating lack of generalisability, without explaining why. Stronger responses would give a description from the study, for example, lack of generalisability due to the study just looking at one type of illness and in one GU clinic in northern UK.

Question 6

(a) This question required candidates to describe what psychologists have discovered about management of stress. Stronger responses had a clear understanding of treatments for stress and included correct studies that support the treatments. For neurochemical benzodiazepines (BZs) and the effect on GABA and how this works to reduce stress, also SSRIs and dopamine, though more explanation was needed to link to stress rather than just anxiety. Also, some clear procedures for biofeedback and imagery with specific links to the studies by Bridge and Budzynski. Stress inoculation was also well detailed in terms of the procedure. The stronger responses covered all three in some detail. Weaker responses often gave limited descriptions with significant errors. 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. There were a lot of Level 2 responses due to the description lacking detail and some adequate use of psychological terminology. Highly detailed descriptions of two or more of the studies can achieve top band.

(b) This question required candidates to evaluate what psychologists have discovered about stress, with the named issue being experiments. Some candidates appeared to struggle with the named issue and gave generic responses. Other evaluation issues were reductionism and application to everyday life and were overall answered very well. Candidates need to link their evaluation points to the studies described in **part (a)**.

There were many Level 1 and 2 responses due to limited analysis. Other responses focused on the practical applications / usefulness. The best responses covered this question by structuring it by issues, along with supporting examples from the studies in **6(a)** and analysis. Candidates should be aware that if the named issue is not addressed, a maximum of 5 marks can be awarded and if only the named issue is addressed, a maximum of 4 marks can be awarded.

Psychology and Organisations

Question 7

(a) Although the question does not require candidates to outline the Fox et al. (1987) study, many did so and achieved full marks. This was answered well with lots of 2-mark responses for a detailed outline of the use of token economy to reduce accidents at work.

(b) Many candidates could describe the study by Oldham and Brass (1979) in some detail, including sample, location, procedure, and results. Weaker responses got details confused in terms of the procedure.

(c)   The most likely strength that candidates used was in-depth data and for the weakness poor generalisability. Top band answers required candidates to show a clear understanding of the question and explain one strength and one weakness with clear links to the study by Oldham and Brass (1979). Many responses would give a weakness and a strength but not in any detail or without any link to the study.

Question 8

(a)   This question required candidates to describe what psychologists have discovered about measuring job satisfaction. The stronger responses covered all three elements listed in the syllabus for this topic in some detail. Weaker responses often gave limited descriptions with significant errors. Some responses gave descriptions of the questionnaires about measuring job satisfaction, with inaccuracies, often getting the three studies muddled with each other. 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.

(b)   This question required candidates to evaluate what psychologists have discovered about measuring job satisfaction, with the named discussion issue being psychometrics. Stronger responses demonstrated clear and detailed understanding of psychometrics, while others appeared to struggle with the named issue. Other evaluation issues used were types of data, usefulness and reductionism and holism. There were many Level 1 and 2 responses due to limited analysis. The strongest responses covered this question by structuring it by issues, along with supporting examples from the studies in **8(a)** and analysis. Candidates should be aware that if the named issue is not addressed, a maximum of 5 marks can be awarded and if only the named issue is addressed, a maximum of 4 marks can be awarded.

PSYCHOLOGY

Paper 9990/32
Specialist Options: Theory 32

Key messages

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

It is important that candidates are made aware of the terminology, theories, disorders and studies identified in the syllabus as some were unable to identify and/or define the terms/concepts 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, technique or 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 results, 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, Anxiety disorders: explanations of phobias: psychoanalytic (Freud, 1909) to help the candidate identify which part of the syllabus the question is referring to as some candidates described the incorrect self-report or theory. For studies, the candidate should learn the aim, sample (sampling method if known), method, procedure, two results (if possible) and conclusion.

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 or method. 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 self-report 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, technique and self-report to help candidates prepare for these questions.

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

This question comes from one of the bullet points in the syllabus. Candidates could describe the three (or four) studies, theories, characteristics/explanations/treatments of disorders or techniques identified in the syllabus 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 response to achieve full marks by describing at least two of the studies, theories, characteristics/explanations/treatments of disorders or techniques and this would need to be a very detailed description. It could be useful for candidates to create revision notes with the title of each bullet point as the header. Alternatively, candidates could create a mind map and put the bullet point in the centre.

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

This question asks the candidate to evaluate the studies, theories, characteristics/explanations/ treatments of disorders 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 strong supporting examples from the studies, theories, characteristics/ explanations/treatments of disorders or 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 to structure the response 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 candidates did not evaluate using the named issue. Quite a few of the answers were structured by study/theory/technique 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

Many candidates appeared to be well prepared for the exam and showed strong knowledge, understanding and evaluation throughout their responses. Some candidates appeared to be less prepared and showed limited knowledge and understanding with brief, superficial and sometimes anecdotal responses. These candidates often had limited evaluation skills.

Time management for this paper was strong for the majority candidates and most attempted all questions that were required. A number of 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 health and organisations.

Comments on specific questions

Psychology and Abnormality

Question 1

(a) There were many strong responses to this question which outlined the characteristics of a specific phobia. Responses were varied with a choice of a wide variety of phobias including button, blood and agoraphobia as popular choices. Popular responses included reference to extreme/high levels of fear, physiological reactions to experiencing the phobic object/situation such as fainting or a panic attack and also avoiding the phobic object/situation. It was common for responses to just state that the person had a 'fear' of the object rather than an extreme fear which was not creditworthy.

(b) Most responses for this question achieved in the 1–2 mark band. Many were able to identify that the horse was considered to represent the father in the little Hans study. Some could identify conflict between id and superego and then explain how this leads to a phobia from the psychoanalytical perspective. There were a few strong responses that were then able to expand their response by explaining how a phobia is a defence mechanism and the phobic object can be a displacement of the object/person that the patient is really frightened of. There were many incorrect explanations. Many stated there was a conflict between the ego and id which was not creditworthy. In addition, many responses outlined the little Hans study in detail without giving any reference to the psychoanalytical explanation of a phobia other than mentioning that the horse represented Hans' father. Many responses outlined that Hans had witnessed an accident involving a horse but this is not a psychoanalytical explanation of phobias and therefore was not creditworthy. Some responses outlined the behaviourist explanation of phobias which was also not creditworthy.

(c) Many strong responses although it was more common to score 4 marks rather than 6 marks as most explained a strength/weakness of case studies and little Hans rather than the psychodynamic explanation of phobias. Common points included qualitative and in-depth detail, lack of

generalisability, bias of the results from the study about little Hans. The strongest responses started with a point about the explanation and then linked it to little Hans as an example before returning to discuss the explanation of phobias at the end of the strength or weakness. One particularly strong answer discussed how the psychodynamic explanation really would not work with some other phobias like the dark or heights because it was hard to see a role for id suppression causing anxiety in these cases.

Many responses did not sufficiently understand what is meant by the psychodynamic explanation of phobias to give responses entirely focused on this. Some responses attempted to explain that the explanation/study was not ethical but this is incorrect as Hans' father had given consent and therefore this type of response was not creditworthy. Some responses attempted to argue that the explanation is reductionist but this is incorrect, unless the response clearly stated that the explanation ignored genetic causes, and therefore was not creditworthy.

Question 2

(a)   Responses varied 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 treatment and management of obsessive-compulsive and related disorders. The strongest responses outlined three of the treatments and included details of the studies by Lovell et al. and Lehmkuhl et al. Other common treatments included outlining how SSRIs can reduce the symptoms of OCD and some included details of research investigating the effectiveness of SSRIs. ECT was referred to in a few responses. This was often less clear as some responses could explain the procedure of having ECT done but were not able to explain how it might reduce the symptoms of OCD.

Weaker responses often gave brief descriptions and inaccurate details of some of the treatments. Inaccuracies were often seen with SSRIs. Many responses gave an incorrect outline of how SSRIs work or no details of how this could lead to a reduction in symptoms. In addition, many candidates gave the side effects of treatments which is not creditworthy in **part (a)** of this question and should be given in **part (b)**. Some candidates gave lengthy descriptions of the symptoms of OCD which was not creditworthy.

(b)   Many of the responses achieved in the Level 1 or Level 2 mark band with a few providing clear analysis and examples from **part (a)** to back up their evaluative points that enabled these type of responses to achieve Level 3 and above. The vast majority of responses covered the named issue of reductionism versus holism. Many were able to correctly explain how the SSRI treatment was reductionist with clear evidence (reference to SSRIs altering the levels of serotonin). Stronger responses evaluated the debate discussing how focusing on the biochemical treatment meant the treatment might not be very effective for those whose OCD is due to cognitive reasons. Likewise, with holism being linked to CBT and ERP, stronger responses correctly identified the reasons why with reference to strengths and weaknesses, such as it helps individual people as their treatment is devised for their specific symptoms and will likely be more effective. Other popular evaluation issues used included determinism versus free-will, nature versus nurture and ethics (with a focus on side effects).

Weaker responses had a tendency to focus on many issues per treatment rather than considering each issue in turn which would have enabled the candidate to provide analysis. These type of responses often defined the issue and debate but not link it to the treatment and/or study or only stating the issue and debate applied to a treatment or study but simply stating that the study and/or treatment supported one side of a debate (e.g. it is reductionist) with little to no explanation for why this was the case. It was common to state that a treatment just ignored the other treatments as the explanation for reductionism. These type of responses often achieved in the Level 1 band.

Psychology and Consumer Behaviour

Question 3

(a)   There were many strong responses to this question. Full mark responses often did a diagram of the 'grid' layout followed by a brief outline of it. Common descriptions included stating that there were long/rectangular aisles that run parallel to each other. 1 mark responses usually outlined or drew the long/rectangular aisles. These responses were frequently brief or did not label their diagram. Some responses just stated that the layout was a 'grid' without further explaining what this means

which was not creditworthy. A significant number of responses gave results from the study by Vrechopoulos which also was not creditworthy.

(b) There were many strong responses to this question. Common details included the independent variable, procedure, method, and results. There were some very strong responses with detail given on the number of participants in the sample as well as the numerical results from the study (although numerical results were not required to achieve marks for the results). Weaker responses often had less detail with some either just outlining the procedure of the study or just giving the results. A few responses stated that the study compared different types of music, which was not creditworthy. In addition, a small number of responses outlined the study by North et al. on music in restaurants which also was not creditworthy.

(c) The vast majority of responses were able to achieve marks for this question and responses achieved marks across the full range of the mark band. Common strengths included ecological validity, generalisability and lack of demand characteristics. Strong responses identified the strength, explained this with an example from the study and frequently stated the effect this would have on validity. Weaker responses were often less well developed and either did not give a clear example from the study by Guéguen et al. or did not explain why the issue raised was a strength. A few responses did explain weaknesses of the study which was not creditworthy.

Question 4

(a) There were some strong responses to this question. Some provided details of the theories about 'buying the product' including the theory of planned behaviour, black box model, consumer decision model. There were some strong descriptions of the theory of planned behaviour that included all elements of it. Black box model sometimes included some clear examples of the 'stimuli', what the consumer may think about when considering a purchase and then the 'response' to either purchase or not. Some of the descriptions of the consumer decision model included all of the steps as well as an explanation of each. Weaker responses often lacked detail of each theory or included some theories from other parts of the syllabus which were not creditworthy as they were not linked in any way to the theory. For example, outlining McCarthy's four Ps for the Black Box Model without indicating that this is the 'stimuli' part of the model. There were a significant number of responses that outlined other parts of the syllabus in their response without giving any details of the three theories in the question. These types of responses were not creditworthy.

(b) If the response outlined creditworthy material in **part (a)** then the vast majority were able to access marks in this part of their answer. Most creditworthy responses addressed the named issue of practical applications. Many were able to give clear suggestions of how these theories could be used by the manufacturer/retailer in order to increase sales. A few did attempt analysis but this was frequently just stating that two theories were similar in that they both had practical applications which is limited analysis. Instead, responses could have explained why the theory might be impractical to the retailer. Other common issues included individual versus situational explanation, cultural bias, reductionism (what was not considered) versus holism (the factors that were considered). Cultural bias was also a popular option but done with more variable success as often the arguments were not very convincing as to why this should apply to a particular model and frequently just identified that the theory was 'western' without explaining why this theory might or might not apply in other cultures.

Psychology and Health

Question 5

(a) Most responses were able to achieve limited credit by giving an example of organic pain such as tissue damage, broken bones, pain from illness, burns, etc. Some responses gained full marked by stating that organic pain is a physical (rather than psychological) pain. A significant number of responses explained what is meant by chronic and/or acute pain which was not creditworthy. Some responses outlined psychological pain and gave phantom limb pain which was also not creditworthy.

(b)(i) There were many strong responses to this question with many giving a clear outline of the Wong-Baker rating scale. Common details included a diagram of some/all of the faces, correct labelling of the faces and/or outline of the 0–10 rating scale. Weaker responses often had errors including the

incorrect range or labels. Some simply stated that it is a measure using faces without any further explanation or detail.

Responses that were not creditworthy included stating it is a measure of pain in children despite this being mentioned in the question or stating that the children colour in the faces which is for the paediatric pain questionnaire.

    (ii) Most responses achieved limited credit for this question on how this scale is used to measure children's pain. Most stated that the children circled or indicated the face that was similar to their pain. A few responses stated that it is used with children over the age of three. A common error was to state it is used with children under three years. Some also indicated that it was a self-report used by health professionals to indicate the level or severity of pain experienced in children. Some responses re-stated what they had said in **part (b)(i)** which was not creditworthy.

(c) There were a number of strong responses to this question and the vast majority were able to give at least one difference between the Wong-Baker scale and the McGill pain questionnaire. Common differences included intensity of pain measured by Wong-Baker versus location of pain measured by McGill, quantitative versus qualitative data collected and simplicity versus complexity of the pain measures. Stronger responses identified the difference and gave an example from each self-report. Weaker responses often gave an example from one of the self-reports or just indicated the difference without any example from either Wong-Baker or McGill. A few of the responses repeated the difference so were not able to gain credit for the second difference or stated that the difference was that one is used with children and one is for adults which was not creditworthy.

Question 6

(a) The responses to this question covered the full range of the mark scheme. Stronger responses gave clear and often detailed descriptions of health promotion in schools, worksites and communities including details of the studies by Tapper et al., Fox et al. and Farquhar et al. Higher mark responses provided clear details about the Tapper et al. study including the age range, procedure, the duration of the study and results. For the Fox et al. study, details were given about the two mines, the use of the stamps and what they could do with the stamps, plus results. For the Farquhar et al. study, strong responses included details on the cities, the use of the surveys and the results. Some included the percentages. Weaker responses when outlining the Farquhar et al. study often did not describe the educational campaign. The responses would give an outline of the cities and the collection of physiological data but it was unclear what the researchers actually did in the experimental cities. A common error was to state that the researcher/employer in the Fox et al. study took away tokens when the worker had an injury which was incorrect. Instead the employee did not gain a token if there was an injury.

There were some responses that were generic. There were some very long responses that included irrelevant, anecdotal description of how schools and worksites might encourage candidates and staff to be healthy. These type of responses were not creditworthy.

(b) Most responses achieved Level 1 or Level 2 for this question. Most did the named issue of the longitudinal research method and could apply this to at least one of the studies described in **part (a)**. Stronger responses were able to give some of the strengths and/or weaknesses of the longitudinal research methods and use the studies as examples. Weaker responses tended to be underdeveloped and often just stated how long each of the studies were without anything evaluative being mentioned in the response. Many indicated that the Fox et al. study was snapshot which was incorrect. Other common evaluation issues were generalisability with stronger responses linking this to cultural bias with clear examples of how participants from another culture may have responded differently to the study. In addition, responses included issues on practical applications, response bias in questionnaires and ethics. Weaker responses often lacked detail with the candidate stating that the study was either positive for the evaluation issue (e.g. it had good ethics) or negative (e.g. it was not generalisable) without much of an explanation as to why this was the case. These type of responses often achieved in the basic mark band.

Psychology and Organisations

Question 7

(a)   There were a variety of responses to this question with many being able to outline what is meant by 'sins of commission' and some responses simply stating it was related to information being inaccurate. Strong, full mark responses frequently identified one of the types such as sunk cost, hindsight bias, belief perseverance or extra-evidentiary bias with a clear definition of what the term means. Limited responses often stated that it is 'misuse of information' with no additional explanation. A few appeared to misunderstand the term and wrote about it being where groups do not agree or groups that just automatically agree with each other. In addition, some responses outlined what is meant by groupthink which was not creditworthy.

(b)   There were many strong responses to this question with many achieving in the 3–4 mark band. Many responses identified that there are nine roles within a team divided into 3 categories. The categories were often identified along with the roles which were frequently defined. In addition, some responses mentioned the Belbin team inventory and/or that roles are determined by the personality and characteristics of each team member which often enabled these responses to achieve full marks. Weaker responses often just identified categories and roles with no clear description given of the theory of team roles. The vast majority of responses had some knowledge of the roles as very few achieved no marks unless they did not attempt the question which was rare.

(c)   Most responses were able to achieve Level 1 or Level 2 for this question. Common strengths included the practical application of the theory with examples of how organisations could use the theory to improve team work, holistic/comprehensive theory and generalisability. Common weaknesses included the difficulty in applying to smaller organisations and the possibility that an employee could partially fit into more than one role.

Stronger responses identified the strength/weakness, explained how it is a strength/weakness and gave an example from Belbin's theory to explain the point. Weaker responses did not always relate back to the roles/categories or theory and therefore achieved fewer marks for this question. Some responses were limited in that they just briefly identified the point (such as stating that it could help to improve team work within an organisation) with no explanation or example to back up their point.

Question 8

(a)   There responses to this question covered the full range of the mark scheme. There were many strong, detailed responses to this question. Strong responses would give a clear description of universalist theories with an outline of the theory that the 'leader is born' with characteristics of leaders including transformational and charismatic leaders. In addition, strong responses gave an outline of the behavioural theory of leadership; leadership qualities can be learnt. Many outlined the types of leadership according to Ohio and Michigan State studies with an explanation of each. Descriptions of adaptive leadership were strong when they gave definitions of each of the 6 principles. For levels of leadership, there were many clear descriptions that stated all three levels with a definition and example as well as stating the importance of personal leadership according to Scouller.

Weaker responses often only outlined the universalist theory and either omitted the behavioural theories or just stated that 'leadership qualities can be learned' but not expanding on the Ohio and Michigan State studies. Adaptive leadership was sometimes described through listing some of the 6 principles without explanation or giving a more anecdotal definition of what it is for a leader to be willing to change depending on the organisational environment/business environment which achieved fewer marks for this question. Finally, weaker responses which outlined the three levels of leadership often just identified the three levels with no explanation.

Some responses described leadership styles from other parts of the syllabus such as contingency theory, situational leadership and/or styles of leader behaviour which was not creditworthy as the three theories were named in the question.

(b)   The marks for this question were commonly in Level 1 and Level 2. Most responses attempted the named issue of determinism versus free-will and some were able to give a clear explanation as to which side of this debate each of the traditional and modern theories of leadership supports. Many

of the responses were very superficial for this issue and merely stated that each theory supported determinism (or free-will) with no explanation given. Other common evaluation issues included nature versus nurture, cultural bias, individual and situational explanations and practical application. A few responses had analysis that extended beyond the limited range by providing an evaluation that explained why theories might be similar or different in terms of the issue being discussed. These responses sometimes achieved analysis through an explanation of why a theory could be considered both deterministic and also free-will or both individual and situational. However, the vast majority of responses were either basic or limited due to the response frequently just identifying a point and then stating that the theory either supported one side of the debate or the other without any explanation or examples to back up these points.

PSYCHOLOGY

Paper 9990/33
Specialist Options: Theory 33

There were too few candidates for a meaningful report to be produced.

PSYCHOLOGY

> **Paper 9990/41**
> **Specialist Options: Application 41**

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

There was evidence to suggest that many candidates had not studied two options with a few candidates answering questions from one option only. Some candidates answered questions from three and even four options. Whilst answers to one option were often very good, some answers to the second option were very 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 marks were limited.

Candidates should not use terms without explanation. Frequently answers stated 'it is reductionist' or 'it is useful in everyday life' without further explanation. To state 'it is reductionist' is merely to identify; it is not automatically a strength or weakness. Further many candidates assume that to be reductionist is always a weakness. It is not; any experiment is reductionist because variables are controlled and only the IV is manipulated. Reductionism is the basis of any experiment and as such it is a strength.

(iv) 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.

(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, 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

Many candidates conduct an experiment whatever the question. An interview, questionnaire or observation are methods independent of an experiment and 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, rather than describe what they know about that topic area, and answers that do not 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.

Section A

Question 1

(a) Many answers were awarded full marks for explaining how SSRIs work when treating OCD (i.e. the question was answered precisely). Some candidates thought that SSRIs block serotonin levels and some candidates focused on the topic area of depression (rather than OCD) where SSRIs are also used.

(b) (i) Most candidates gave an appropriate explanation of what is meant by an obsession, a recurring and persistent thought that interferes with normal behaviour, and were awarded 1 mark. Most candidates were able to give an appropriate example such as a fear of contamination, and were awarded the second available mark. Some answers confused obsessions with compulsions and some used examples from impulse control disorders which were incorrect.

(ii) For this question part most candidates gave an appropriate explanation of what is meant by a compulsion and were awarded 1 mark. Examples such as double checking, excessive washing or cleaning or accumulating things (hoarding) were all correct and were awarded a further mark. A number of candidates wrote about body dysmorphic disorder (BDD) which also received credit. Candidates writing about impulse control disorders received no credit.

(c) Many candidates described in detail the case study of Jason, outlined by Lehmkuhl et al. (2008), often scoring maximum marks for the inclusion of specific detail related to his contamination fears

such as doorknobs and library books. Such answers also described his exposure and response prevention (ERP) treatment allowing Jason to touch papers for example. A number of candidates described the case study of Charles, but while Charles had OCD he was not treated with ERP. Some candidates described the study Saavedra and Silverman about the boy who was treated with ERP but who did not have OCD.

(d) Answers to this question part followed the same pattern as **part (d)s** for all other options, which was a lack of relating strengths and weaknesses to the question set. In this instance drug treatments had to be related to obsessive compulsive disorder. A top mark answer typically included: 'drugs are easy to take, nothing more than swallowing a pill, such as swallowing an SSRI pill for OCD'. In this answer there is a strength and that strength is related to the question. If this process is repeated for another strength plus example and two weaknesses plus examples, then full marks will be achieved.

Question 2

(a) Some candidates provided excellent answers explaining that choice blindness is a failure to recall a choice immediately after we have made that choice. Quite a number of candidates did not appear to know the term 'choice blindness', which is a fundamental component of the study by Hall et al. (2010). Nearly all guesses were incorrect.

(b) A common error in response to this question was to provide a generic strength of a field experiment but then not go on to provide an example from the study, which the question required. For example, a candidate might write 'the setting is high in ecological validity', which would be awarded 1 mark, but not go on to say that 'the Hall et al. study was conducted at a stall in a supermarket', when doing this would allow the second mark to be awarded. Candidates should ensure they address the whole question if they wish to receive the full allocation of marks.

(c) All candidates could describe deception, but some struggled to relate two instances of deception in the Hall et al. study. Some candidates gave incorrect examples and were awarded no marks, others were awarded 1 mark for stating, for example, 'they were deceived about the aim of the study' but needed elaboration for further credit. Answers achieving higher marks often extended the deception regarding the aim and quoted from the study itself when writing 'we recruited the participants by asking them whether they were willing to take part in the quality control test of the jam and tea assortment at the store'.

(d) Many answers included two strengths and two weaknesses of using an opportunity sample, but often only scored partial marks because frequently answers were not related to the study on choice blindness as the question required. A strength such as 'people who are 'readily available' can be chosen to participate' (1 mark) could extend to 'when they are passing a stall located in a supermarket' would have been sufficient for 2 marks. Many answers had no appropriate conclusion.

Question 3

(a) Many candidates could explain what an independent measures design involves, although some candidates wrote incorrectly about a repeated measures design. Some candidates explained that the design involved different communities (cities) and top mark answers explained that there were two independent treatment communities compared with three reference or control communities.

(b) A number of candidates appeared to be confused about the aim of this study, suggesting that cities were chosen because of high rates of cardiovascular disease. The point of the study was to prevent the onset of cardiovascular disease, not to treat it. This meant that answers suggesting cardiovascular disease was one of the criteria for selection scored no marks. Some candidates incorrectly suggested age was a factor but this is also incorrect. Two correct answers where that, for example, the population of the community needed to exceed 30,000 people, and that the communities were located in Northern California.

(c)(i) Some candidates incorrectly defined the term random sample. Other candidates correctly wrote that it might involve a computer selecting a random sample or by putting every name into a hat. However, the question stated 'used in this study' and so there had to be a link to the study. Some candidates did not do this but others suggested using electoral registers or census details to determine which names would go into the hat or computer.

(ii) Candidates often gave a general answer without relating it to the study in question. This meant that answers such as 'so every person has an equal chance of participating', or that 'it eliminates researcher bias' could only be awarded 1 partial answer mark. Candidates scoring 2 marks gave some elaboration that was clearly related to the study such as 'it eliminates bias of choosing people from one specific region which may or may not be healthier than people from another'.

(d) Candidates did not always achieve marks because they did not relate strengths and weaknesses of physiological measures to community health promotion projects as the question required. Many candidates summarised what they had already written, rather than provide a conclusion and so no mark could be awarded. Reference to the mark scheme shows that a conclusion is defined as a 'decision reached by reasoning' and this is what candidates should do.

Question 4

(a) A large number of answers were a commonsense statement such as 'a reward without money' without further elaboration. Candidates need to providing answers in sufficient detail to show their knowledge and understanding. To go beyond this statement candidates could have referred to non-monetary rewards being intrinsic rather than extrinsic or they could have given an example such as praise or recognition. Doing either of these would show appropriate knowledge understanding of non-monetary reward systems.

(b) Many candidates scored maximum marks for identifying and showing understanding when explaining two types of non-monetary reward. Such rewards include praise, respect, recognition, empowerment and sense of belonging.

(c) Candidates had to suggest how non-monetary rewards apply to two of Maslow's needs. Many answers cited the need for self-actualisation which can result from rewards such as praise, respect and recognition. The need for self-esteem was commonly used and again this could result from praise or respect or recognition for high quality work. Some candidates chose the need of love and belongingness (social need) but did not always relate this to organisations and examples could only be credited if it was related the need for affiliation at work with work colleagues.

(d) Candidates often gave short answers and it was unclear what was a strength or weakness or link to motivation. For example, a candidate might write 'workers are motivated by money' without further elaboration. Some candidates continued the theme of non-monetary rewards, but this question part focused on monetary rewards.

Section B

Question 5

(a) Candidates scoring high marks sampled men and women who had undergone face-to face-therapy for a period of time and then interviewed them over the phone asking them questions about how the face-to-face therapy was for them. Evidence was often gathered using both closed and open questions. Many candidates compared face-to-face therapy with telephone therapy and did not investigate the effectiveness for women compared to the effectiveness for men, so did not address the question set and marks were limited.

(b) The psychological evidence quoted by most candidates was a study by Lovell et al. (2006) which concluded the therapy delivered by telephone was as effective as face-to-face therapy. What most candidates did not do was link this study to their design as a question required. Often methodological decisions were inappropriate for candidates designing an incorrect study. Further, many candidates evaluated their design decisions but the question does not ask for evaluation.

Question 6

(a)   This question on 'colour choice' appeared to be a popular choice but there were anecdotal common-sense answers which could not be credited. Candidates could have been awarded some marks if they had used appropriate knowledge about interviews and the type of questions that would have been appropriate to the requirements of the study. Answers at the top of the mark range included the type of interview (such as face-to-face), the setting (conducted in a supermarket), sample questions (most often closed) and how these questions would be answered/scored (often rating scales to gather quantitative data).

(b)   Very few candidates wrote about the work of Grossman and Wisenblit (1999), who studied product colour and associative learning, for their psychological knowledge. Instead answers were often anecdotal with suggestions like 'I included the colour green in my interview because vegetables are green'. Many candidates incorrectly evaluated their designs.

Question 7

(a)   Some candidates provided excellent answers when designing an experiment to address the question set, often allocating participants to two conditions of the IV (those using pill counting and those using biochemical tests). An independent measures design was apposite although some candidates suggested a repeated measures design which was problematic because a biochemical test would be irrelevant if patients were not taking pills. Top mark answers applied an appropriate range of specific experimental features whereas middle and bottom band answers did not, and were too 'story like'.

(b)   Some candidates wrote extensively about the work on pill counting by (Chung and Naya, 2000) and on biochemical tests by (Roth and Caron, 1978). This was often substantial but according to the mark scheme such description could only be awarded a maximum of two marks. The second two marks are awarded for describing how this psychological knowledge informs the design in **part (a)** and often this second component was not done at all by many candidates. For methodological evidence, a significant number of candidates evaluated their designs. Candidates should explain their design decisions, such as why they used an independent measures design rather than a repeated measures design.

Question 8

(a)   Most candidates used a questionnaire as the main method, as the question required, but often methodological knowledge about questionnaire design was limited and often questions did not investigate cognitive limitation/error. At this level it is expected the candidates know about question technique and format. Questionnaire technique is whether the questions are done online or using 'paper and pencil' in front of an examiner. Questionnaire format is the use of open or closed questions (or both) and this would include examples of questions, how the questions are scored and how they will be interpreted by researchers. A second problem was that candidates often did not appear to know about the different types of cognitive error such as those outlined by Forsyth (2006) and this meant that designs were often very general or superficial rather than being based on psychological knowledge.

(b)   Following on from **(a)** above, psychological knowledge was in many instances very limited or absent. Some candidates did write about Forsyth's sins of commission, sins of omission or sins of imprecision, and in top mark answers these types of error informed the questions participants were asked, i.e. these candidates were using relevant psychological knowledge to inform their design. Methodologically, as also mentioned in **part (a),** knowledge of questionnaires was limited in some answers but was very strong in others by inclusion of many of the features of questionnaires mentioned above.

Section C

Question 9

Answers scoring the highest marks were those which were organised and answered the question set. Organised answers frequently included two debates: whether the use of drugs was effective for *all* patients or not (and this often included the consideration that it might be effective for some people but not others); secondly whether the use of drugs was preferable to alternative forms of treatment for schizophrenia such as electroconvulsive therapy or cognitive behaviour therapy. Answers at the bottom end of the mark range often described the history of drug treatments, mentioning first, second and third generations. Other answers at the bottom of the mark range simply gave generic strengths and weaknesses of taking drugs with no attempt to relate these to schizophrenia.

Question 10

Many candidates described the choice heuristics of availability unrepresentativeness and sometimes these were related to decisions about purchasing goods and sometimes they were not. There are no marks for description in **Section C** questions. Descriptions can be included as part of evaluative discussion and appropriate credit would then be awarded. Candidates rarely contrasted heuristics with alternatives. For example, Knutson et al. studied pre-cognitive decisions where the brain has already made a decision before a person can make a cognitive decision. Similarly candidates did not always consider that the anchors of availability and representativeness might apply less compared to other anchors/heuristics.

Question 11

There were many very strong answers which considered the arguments both for fear arousal as a strategy for promoting health and the arguments against the use of fear arousal. The research study of Janis and Feshbach (1953) was often used to support the arguments being made. These answers also considered alternative strategies such as providing information and brought in the work of Lewin (1992). Occasionally these top band answers also included the work of Cowpe (1989) who used a combination of both fear arousal and providing information. Answers achieving middle band marks considered nothing more than the strengths and weaknesses of fear arousal and answers at the bottom band merely described the study by Janis and Feshbach or anecdotally considered whether it is good or not to scare people into improving their health.

Question 12

The most logical approach to achieve a top mark was to consider how setting goals would improve motivation and include a debate about whether setting 'smart' targets would increase motivation or not. This could be followed by bringing in alternatives to need theories. For example, goal setting would be irrelevant for McClelland's need for power or affiliation. Whilst this approach was taken by a small number of candidates most merely wrote about Maslow's hierarchy of needs and did not consider goal setting as proposed by Latham and Locke (1994).

PSYCHOLOGY

Paper 9990/42
Specialist Options: Application 42

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

There was evidence to suggest that many candidates had not studied two options with a few candidates answering questions from one option only. Some candidates answered questions from three and even four options. Whilst answers to one option were often very good, some answers to the second option were very 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 marks were limited.

Candidates should not use terms without explanation. Frequently answers stated 'it is reductionist' or 'it is useful in everyday life' without further explanation. To state 'it is reductionist' is merely to identify; it is not automatically a strength or weakness. Further many candidates assume that to be reductionist is always a weakness. It is not; any experiment is reductionist because variables are controlled and only the IV is manipulated. Reductionism is the basis of any experiment and as such it is a strength.

(iv) 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.

(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, 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

Many candidates conduct an experiment whatever the question. An interview, questionnaire or observation are methods independent of an experiment and 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, rather than describe what they know about that topic area, and answers that do not 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.

Section A

Question 1

(a) Many answers were awarded full marks for outlining how operant conditioning explains token economy. Some candidates confused operant conditioning with classical conditioning and some answers described token economy with no reference to operant conditioning.

(b) Most candidates were able to outline the findings of the Paul and Lentz study, often in enough detail, or by including a range of different features, to be awarded full marks.

(c) Many answers to this question were very strong. Most answers referred to problems of cross-over effects where one treatment would interfere with another if a repeated measures design was used and it would not be known which technique resulted in reduction of their schizophrenia. Some candidates referred to an independent design reducing demand characteristics. However, participants know they have schizophrenia and they know what their treatment is. This means that they know the 'aim' of the study and demand characteristics do not apply. Notably the participants in the Paul and Lentz study were real people with a real illness.

(d) Answers to this question needed to relate the strengths and weaknesses to the question set. In this instance answers had to be related drug treatments for schizophrenia. A top mark answer typically included: 'drugs are easy to take, nothing more than swallowing a pill, such as swallowing an atypical anti-psychotic pill for schizophrenia'.

Question 2

(a) Most candidates could be awarded full marks for full and accurate answers. A few candidates suggested there was no harm, and although there was no harm, this could not be supported with any evidence from the study.

(b) Some candidates explained two of round trip, central trip or wave trip and were awarded full marks. Other candidates could identify two of these, but needed to provide more detail for higher credit. Some candidates wrote about shopper behaviour patterns (i.e. tourist, explorer raider, etc.) rather than types of movement pattern.

(c) (i) In this instance the study involved shoppers in a supermarket and this meant that one feature of the interview was that it was face-to-face. A second feature of the interview was that the interviews were structured (semi structured was also credited) with a predetermined set of questions that all participants were asked. Some candidates wrote generically about interviews, sometimes scoring marks, but more frequently 0 marks were awarded (for example stating that the interviews were unstructured).

(ii) If candidates had suggested a structured interview in **part (c)(i)** a logical strength was that all participants were asked the same questions in the same order and so direct comparisons of different types of shopper movement patterns could be made. Another strength is that a face-to-face interview is more personal/friendly than other types of interview and non-verbal aspects of the conversation can be noted. A face-to-face interview is also more likely to continue for longer than a telephone interview.

(d) Many answers included two strengths and two weaknesses of using CCTV to investigate shopper behaviour, but often only scored partial marks because frequently answers were not related to the specific question set. A weakness such as 'a CCTV camera has a limited panorama' was a common limited response which could have been extended to 'so some shopper behaviours may be missed or interpreted' which would have been sufficient for 2 marks. Many answers needed to include an appropriate conclusion.

Question 3

(a) Many candidates struggled to provide an answer, or appeared to guess incorrectly. A few candidates described the GAS model but this did not answer the question set.

(b) There are generally 3 features of a Type A personality which include: being competitive, being time conscious, and being aggressive and assertive. Identification of any two of these three achieved limited credit and elaboration of two of these could result in being awarded full marks available. A number of candidates referred to other features of Type A personality and these were also given appropriate credit.

(c) (i) Many candidates were able to suggest two appropriate causes of stress, scoring 1 mark for identifying life events and a further 1 mark for describing the features outlined by Holmes and Rahe (1967). The second suggestion was often 'work' and elaboration usually came from the study by Chandola et al. (2008). Many candidates suggested that studying A levels or having relationships is stressful, but without supporting psychological knowledge these answers are anecdotal.

(d) Candidates did not always relate strengths and weaknesses of longitudinal studies to stress as the question required. Answers were frequently awarded limited credit for generic points 'longitudinal studies often suffer from attrition' but needed to focus on the specific question. Candidates did not always provide a conclusion, or provided a summary instead. Reference to the mark scheme shows that a conclusion is defined as a 'decision reached by reasoning'

Question 4

(a) Most candidates could not explain the term 'job involvement' which is the extent to which the individual identifies psychologically with his/her/their job. Some answers incorrectly referred to organisational commitment, job enlargement or job rotation.

(b) (i) and (ii) Answers to these two question parts resulted from considering the content in fig. 4.1. For **(i)** the correct answer was low job involvement and low organisational commitment. Elaboration was needed of these terms in order to achieve the further 1 mark, or alternatively to provide an outline of what Blau and Boal identified as apathetic employees. For **(ii)** the correct answer was low job involvement and high organisational commitment. Again, elaboration was needed in order to achieve the further mark or to outline what Blau and Boal described as corporate citizens.

(c) This question focused on the four categories of absence identified by Blau and Boal: medical, career-enhancing, normative and calculative. Many candidates outlined two of these clearly and in detail, and so received full marks. Some candidates identified two types as being voluntary and involuntary. Although these were not terms used by Blau and Boal, candidates were awarded partial marks. A few candidates provided anecdotal answers such as 'one category is being ill' which could not be credited.

(d) Only a small number of candidates achieved full marks. It is essential that candidates relate the generic point to the question specifically, in this instance absenteeism. Summaries were often provided rather than conclusions.

Section B

Question 5

(a) A significant number of candidates answered this question incorrectly. Many candidates tried to design an experiment but then struggled to identify an independent and dependent variable, or an experimental design. The use of the experimental method was irrelevant to this study using an observation. Many candidates used virtual reality. However, this would not allow the frequency of characteristics of schizophrenia to be observed. Indeed, many designs required participants to answer questions which again were irrelevant in the design of a study using an observation. Candidates can only be awarded marks if their answer uses the method stated in the question. Candidates scoring maximum marks for this question usually had nurses covertly observing different patient behaviours recorded in behavioural categories over a period of time and then a chart constructed to see which schizophrenic behaviours were the most common.

(b) The psychological evidence quoted by most candidates was a study by Freeman et al. which used virtual reality to investigate persecutory ideation with non-schizophrenic patients. However, this was only peripherally relevant to this question. Directly relevant were the 'standard' schizophrenic behaviours, such as catatonia, loss of speech, lack of typical facial expressions, etc. A common missed point was that participants must have been people with schizophrenia (people with psychosis) and not non-schizophrenic volunteers such as students taking part in an experiment.

Question 6

(a) Most candidates opting for this question chose to design a laboratory experiment, which was a logical choice. Most designs were appropriate and many candidates achieved high marks. However many did not, because many candidates did not demonstrate understanding of what a 'jingle slogan' was, because their designs involved presenting slogans to participants with music playing in the background (or not). Other candidates needed to include elaboration on what the slogans involved or how they were operationalized. Designs in this question part should be based on appropriate psychological knowledge such as the work on 'effective slogans' by Kohli et al. (2007).

(b) Designs in **part (a)** should be based on relevant psychological knowledge and in this instance the work of Kohli et al. was apposite. Some candidates wrote about 'music in restaurants' quoting the North et al. study, which is from a different topic area, jingle slogans did not appear to be well understood.

Question 7

(a) This question was answered incorrectly by many candidates. The focus of this question was stress-prevention techniques (and whether they are more effective for women than men). The question was not about treatments for stress. This meant that although the design of the study, often an experiment, could be credited, when it became 'treatments' no marks could be awarded. Some candidates compared treatments versus non-treatments and did not address the women versus men comparison required by the question.

(b) Many candidates wrote about biofeedback as studied by Budzynski et al (1969) or about imagery (Bridge, et al., 1988). Both of these studies were irrelevant because they focus on treating people who are already stressed. Relevant psychological knowledge was preventing stress as studied by Meichenbaum (1985). For methodological evidence a significant number of candidates evaluated their designs. Candidates should explain their design decisions such as why they used an independent measures design rather than a repeated measures design.

Question 8

(a) Most candidates used an experiment as the main method, as the question required, but often methodological knowledge about questionnaire design was limited and often questions did not investigate which type of monetary reward was the more effective motivator. Some answers were very strong and the full range of specific experimental techniques were applied such as IV, DV, controls, design and type of experiment. Many answers used a questionnaire to gather data from their participants and this is an acceptable strategy.

(b) Psychological knowledge was often strong, but more detail about bonuses and performance related pay could have been included. In top mark answers, performance-related pay was clearly applied, such as 'for every sale over a weekly target of 100 you will receive a 10 per cent bonus'. Methodologically, the use questionnaires was sometimes very strong but often questions were limited to 'Do you prefer bonuses or performance related pay?'. Candidates often evaluated what they had done rather than providing reasons for what they had done.

Section C

Question 9

Answers scoring the highest marks were those which were organised and which focused on the question set. Organised answers frequently included a debate about whether impulse control disorders are caused by high levels of dopamine or not; and if not then a discussion about what does cause impulse control disorders followed with candidates bringing in details of the behavioural approach or the work by Miller. Some candidates described rather than discussed the dopamine hypothesis followed by descriptions of the behavioural approach and Miller's feeling-state theory. Answers at the bottom end of the mark range candidates often described different impulse control disorders such as pyromania, etc. Other answers at the bottom of the mark range focused on the ways in which impulse control disorders could be treated rather than focusing on causes. There are no marks for pure description in any *Section C* answer.

Question 10

Many candidates divided their answers into four parts: a consideration of quantitative data, a consideration of qualitative data, knowledge about brand recognition and the use of children in psychological studies. Whilst this gave candidates a lot to write about, answers were often disjointed because these components were unrelated to any of the others. For example, a strength of quantitative data was given in isolation with no reference to brand recognition, or the ethics of children in studies was described with no reference to brand recognition or types of data. Candidates are strongly encouraged to construct a coherent answer that focuses on the question set.

Question 11

There were many very strong evaluative answers which focused on the question. However, there were many candidates who achieved no more than bottom-band marks because they described the study by Sherman et al. on repeat prescriptions and then described alternative measures of adherence such as using pill counting, biochemical tests, etc. Candidates should be aware there are no marks for description in

Section C questions. Descriptions should only be included as part of evaluative discussion which can then receive appropriate credit.

Question 12

There were a few good answers, but most received limited credit because they did not answer the question set. A typical answer would give a generic response regarding strengths and weaknesses of open-ended questions, but did relate them to job satisfaction. The most logical approach to achieving a top mark would be to consider how open-ended questions could be used to measure job satisfaction but then to debate about ways in which job satisfaction is more usually measured - closed questions or rating scales. This would bring in the JDI (Job Descriptive Index), the MSQ (Minnesota Satisfaction Questionnaire) and Walton's QWL (quality of work life) and allow candidates to show their psychological knowledge of job satisfaction.

PSYCHOLOGY

Paper 9990/43
Specialist Options: Application 43

There were too few candidates for a meaningful report to be produced.