

PSYCHOLOGY

Paper 9990/11 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 the responses. The essay (final question) requires four evaluation points to be present in depth (two strengths and two weaknesses) with at least one of these about the named issue. Credit is limited if the named issue is omitted or just described.

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

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

Candidates need to know about real-world applications for all core studies. To show understanding, answers need to tell the Examiner what the application is based on in the particular core study and then how this could be achieved. Again, this must be explicitly made clear by the candidate.

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

Candidates also need to know the set procedure of studies in the order presented in the original journal article. Questions can be based around just part of a procedure and the candidate must be able to produce an answer that is directed and concise rather than writing about the whole of the procedure. This can sometimes mean a candidate may run out of time for other questions.

There is enough time for answers to be planned to ensure that the response given by a candidate is focused on the demands of each question. This is a crucial skill to develop as some candidates appear to have good knowledge of a study but do not apply this effectively to the question(s) set.

General comments

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

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

There were also some 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 stated the number of participants in each of the experimental groups in the study by Bandura. Common incorrect responses included the total number of participants used in the study, or how many of each sex participated. It is important to carefully read what the question is asking for.
- (b) Stronger responses could clearly outline how participants were allocated to one of the conditions. Popular choices included the aggression scales used by the teacher/experimenter and that they were assigned at random after this observation phase. Common incorrect responses included sex distribution across conditions or a description of the final observation stage of the study.
- (c) There were many correct responses to this question. Common choices included 'sock him' and 'Pow'. There was a significant proportion of responses that gave examples of non-imitative verbal aggression or physical aggression. It is important for candidates to carefully read what the question is asking for.

Question 2

- (a) There were few correct responses to this question. Some candidates could name both ways that each scene had been rated. However, a majority of responses were incorrect, including how the participant rated the scenes or 'emotional intensity' which is what the participant rated the scenes on. This was not what the question was asking. There was a significant number of blank responses to this question.
- (b) There were some clear, concise responses to this question. For example, noting the potential issue of generalisability due to only using females was a popular choice and usually described well. Other popular choices included the lack of mundane realism and the use of self-reports to measure emotion reactions. There was a number of responses that did not provide an example from the study to elaborate on a chosen weakness. Candidates need to be aware that 'in this study' in the question requires an explicit example from that study to be able to access all available marks.

Question 3

- (a) There were many correct responses to this question. Popular descriptions included the sample size, where they were studying, and that they received course credit. However, there were a significant minority of responses that could not describe three features/characteristics of the sample or confused the sample with that from Experiment 2 or a different Core Study. It is important for candidates to note the number of marks assigned to a question as this typically represents (in short answer questions) the number of correct elements that need to feature in a response.
- (b) A small minority of responses could describe a result from the Memory or Belief? Questionnaire. The most popular result was more participants in the Love Group having a memory of loving asparagus compared to the Control Group. With results, where possible, there must be a meaningful comparison to be able to access the marks available. Therefore, simply stating more people in the Love Group had a memory would only receive partial credit. Several incorrect responses chose a different result, or a result based on what the Memory or Belief? Questionnaire was then used for, for example, results about believers and non-believers. Again, it is important for candidates to read the question carefully. There was a significant number of blank responses to this question.

Question 4

- (a) Stronger responses could clearly describe four elements of the experimenter feedback. Popular elements included the use of prods, how it was scripted, and the tone used by the experimenter. Incorrect responses tended to focus on the responses from Mr. Wallace or the procedure of allocating lots. Some responses simply outlined another part of the procedure in the study by Milgram or the findings.

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

Question 5

- (a) The majority of responses could outline the aim of the study by Baron-Cohen et al. Popular choices included to investigate Theory of Mind and to test the newer version of the Eyes Test. There were some responses that only gave a very brief aim which could only be awarded partial credit. Some responses presented a finding from the study so could not gain any credit. It is crucial for candidates to read questions carefully.
- (b) Stronger responses could clearly explain one reason for standardisation. The most popular explanation focused on reliability and replicability with an example from the study by Baron-Cohen et al. Some responses were awarded partial credit for explaining a reason but not providing the final 'in this study' part of the question. Some responses focused on control and cause and effect and gained appropriate credit. The question was why and not how so responses that simply provided examples of how the study was standardised could not be awarded credit.

Question 6

Stronger responses could clearly outline the individual-situational debate and provide clear examples from the study by Pepperberg. Popular examples included how Alex may have been unique with his talents (individual) and how Alex learned quickly using the Model-Rival Technique (situational). There was a significant number of responses that were tautological and could not access marks. For example, stating that the situational side of the debate is about the situation cannot be credited as it is simply re-using the words in the question and does not explicitly show understanding. To improve, candidates should have examples from each Core Study that appropriately support each of the issues and debates at AS-Level. This question had the most blank responses on this paper.

Question 7

There were some strong responses to this question. Some candidates could clearly describe the procedure for this specific condition. Popular elements included the opaque wall, the small window to pass tools through and that the chimpanzees were in adjacent booths. Some responses provided information about a different part of the procedure or did not provide any description of the 'Cannot See' part of the procedure so could only gain partial credit. Candidates need to know the procedure for all Core Studies.

Question 8

- (a) Stronger responses could clearly outline how the score was calculated. However, there was a significant number of responses that misinterpreted the question and provided information about other measures or gave descriptions of false alarms and/or listed the names of party goers. It is important for candidates to be familiar with how measurements were calculated and reported.
- (b) There were some strong responses provided here that fully engaged with the stimulus material. Jamie was the most popular choice for the debate by a significant margin. Popular arguments supporting Jamie included the sample sizes, the gender imbalance and that they were all from a group of willing participants. Popular arguments supporting Seth tended to focus on a wide age range and that the sample size was quite large. Some responses simply described aspects of generalisability without engaging in the Seth-Jamie debate so could only gain partial credit. To improve, candidates need to clearly choose one side of the argument and then explain why they support it using evidence from the study.

Question 9

- (a) For these types of questions, responses should focus on the general psychology that is being investigated in the study rather than a specific aim of the study. Therefore, aspects of the Saavedra and Silverman study that could gain credit here included a definition of a phobia, classical conditioning, evaluative learning, and operant conditioning. Credit could be given to generic descriptions of the principles for all of these. However, many responses focused too narrowly on the aims of Saavedra and Silverman and what was achieved by the therapy. These could only gain one available example mark.
- (b) The responses to this question varied a lot. Strong responses could clearly show understanding of confidentiality using an example from the study by Saavedra and Silverman. One example of a popular explanation was that we only knew his gender and ethnicity. Some responses mis-read this part of the question and wrote about the boy increasing in confidence which is not what the question was about. For informed consent, many responses could explain that the mother and child gave consent after knowing what was going to happen during the therapy. A significant number of responses reported that the child did not give consent, but this is incorrect. The guideline of privacy was known by a minority of responses as many tended to be incorrect in their knowledge about this guideline. Many responses stated that this was about confidentiality (incorrect) or that the therapy was conducted privately (also incorrect). Correct responses could explain how certain information that the boy may not want to tell was needed to be able to progress with the therapy. Protection from psychological harm was a stronger guideline in terms of correct responses. The fact that the boy had to face his fears directly was a popular argument for this guideline being broken. To improve, candidates need to be explicit with their knowledge of an ethical guideline and then provide an example from the study that shows it was broken or not broken. Each guideline is effectively marked out of two.

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 qualitative data. Common choices included 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 that the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these types of questions as it is testing their evaluation skills only. In addition, some responses followed a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics) and some responses appeared to be prepared essays for Dement and Kleitman without one of their points being about qualitative data. A response that fails to have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses used qualitative data in an evaluative sense but did not fully explain why it could be a strength and/or a weakness. Several responses did not cover the named issue, only describing what qualitative data were collected in the study by Dement and Kleitman. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples from the study to show clear understanding. These are the requirements for a Level 4 response.

PSYCHOLOGY

Paper 9990/12 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 present in depth (two strengths and two weaknesses) with at least one of these about the named issue. Credit is limited if the named issue is omitted or just described.

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

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

Candidates need to know about real-world applications for all core studies. To show understanding, answers need to tell the Examiner what the application is based on in the particular core study and then how this could be achieved. Again, this must be explicitly made clear by the candidate.

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

Candidates also need to know the set procedure of studies in the order presented in the original journal article. Questions can be based around just part of a procedure and the candidate must be able to produce an answer that is directed and concise rather than writing about the whole of the procedure. This can sometimes mean a candidate may run out of time for other questions.

There is enough time for answers to be planned to ensure that the response given by a candidate is focused on the demands of each question. This is a crucial skill to develop as some candidates appear to have good knowledge of a study but do not apply this effectively to the question(s) set.

General comments

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

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

There were some 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 majority of responses correctly stated the approximate number of participants used in the study by Piliavin et al. Common incorrect responses included a sample size from a different Core Study.
- (b) Only a small minority of responses could state how the condition was chosen. Most of the responses claimed it was 'random' which is incorrect. Candidates need to know the exact methodology used in all 12 Core Studies. There was a significant number of blank responses for this question.
- (c) There were some correct responses to this question. Common choices included location and speed of helping. There was a significant proportion of responses that gave differences of the victims (e.g., race and ill/drunken). It is important for candidates to carefully read what the question is asking for.

Question 2

- (a) There were few correct responses to this question. Some candidates could name two items that Alex had available to him on request, with water and seed mix being popular choices. However, many responses were incorrect including asking for toys or stating what was same or different, which was not what the question was asking. Candidates need to carefully read what the question is asking for. There were many blank responses to this question.
- (b) There were some clear, concise responses to this question. For example, noting the potential issue of generalisability due to only using one parrot that might be unique. Another popular choice was a description about the lack of mundane realism. There was a number of responses that did not provide an example from the study to elaborate on a chosen weakness. Candidates need to be aware that 'in this study' in the question requires an explicit example from that study to be able to access all available marks. Also, there was a significant number of responses describing an ethical weakness or arguing that it was cruel to keep a parrot in captivity. Neither of these types could not gain credit here as they were not answering the question set.

Question 3

- (a) There were many correct responses to this question. Popular descriptions included the sample size, how many males and female participated and how many were studied intensively. However, there were a significant minority of responses that could not describe three features/characteristics of the sample, or confused the sample with a different Core Study. It is important for candidates to note the number of marks assigned to a question as this typically represents (in short answer questions) the number of correct elements that need to feature in a response, in this case, three.
- (b) A small minority of responses could describe a result from REM awakenings. These responses could describe the difference, and which sample the result had come from (e.g., the five participants studied intensively). With results, where possible, there must be a meaningful comparison to be able to access the marks available. Therefore, simply stating there were more dreams in the REM awakenings could only be awarded partial credit. Several incorrect responses chose a different result to describe (e.g., the number of dreams recalled overall rather than first half compared to second half) or a result based the descriptions of dreams given by participants. Again, it is important for candidates to read the question carefully.

Question 4

- (a) Stronger responses could clearly describe four elements of the interview and dehoax part of the study by Milgram. Popular elements included the debrief, the ratings given on how painful they thought the shocks had been, and that they got to meet Mr. Wallace to ensure he was fine. Incorrect responses tended to focus on the procedure for giving the shocks before the interview and dehoax. Some responses simply outlined another part of the procedure in the study by Milgram or the findings. There was a significant number of blank responses to this question.

- (b) For this type of question, responses must contain two parts. The first is the 'what' – what real-world application could be. The second is the 'how' – how will the real-world application be achieved in an ethical way. Common responses included utilising uniforms or some form of verbal prods to encourage the workers to be more obedient. There were some unethical responses about physical punishment and reducing wages, and these were not given any credit.

Question 5

- (a) The majority of responses could outline the aim of the study by Canli et al. The most popular choice was to investigate the role of the amygdala in memory and emotion. There were some responses that only gave a very brief aim which could only be awarded partial credit. Some responses presented a finding from the study so could not gain any credit. It is crucial for candidates to read questions carefully.
- (b) Stronger responses could clearly explain one reason for standardisation. The most popular explanation focused on reliability and replicability with an example from the study by Canli et al. Some responses were awarded partial credit for explaining a reason but not providing the final 'in this study' part of the question. Some responses focused on control and cause and effect and gained appropriate credit. The question was *why* and not *how* so responses that simply provided examples of *how* the study was standardised could not be awarded credit.

Question 6

Stronger responses could clearly outline what is meant by the learning approach and provide clear examples from the study by Saavedra and Silverman. Popular descriptions included classical conditioning and operant conditioning. Popular choices for examples from the study included being hugged by mum during the therapeutic process, and how he associated the classroom incident with his fear/disgust of buttons. To improve, candidates need to have examples from each Core Study that appropriately support each of the approaches at AS-Level.

Question 7

There were some strong responses to this question. Some candidates could clearly describe the Restaurant Questionnaire. Popular elements included the number of dishes, that it was formatted to look like a menu, and the 1–8 rating scale. Some responses provided information about a different questionnaire or based the responses around willingness to eat, which was not the intention of this questionnaire. Candidates need to know the measurement tools used in all Core Studies.

Question 8

- (a) Stronger responses could clearly outline the category of initiates new activity. However, there was a significant number of responses that misinterpreted the question and provided information about copying the model, providing examples of behaviours that showed this. These responses could not gain credit. It is important for candidates to be familiar with how behavioural categories are defined and reported. This question had the most blank responses on this paper.
- (b) There were some strong responses to this question that fully engaged with the stimulus material. Maryam was the most popular choice for the debate by a significant margin. Popular arguments supporting Maryam included the standardisation of procedure, and inter-observer reliability. Popular arguments supporting Zara tended to focus on that the stooge could act differently for each participant, and variation in participant reactivity. Some responses simply described aspects of reliability without engaging in the Maryam-Zara debate so could only gain partial credit. To improve, candidates should clearly choose one side of the argument and then explain why they support it using evidence from the study.

Question 9

- (a) For these types of questions, responses should focus on the general psychology that is being investigated in the study rather than a specific aim of the study. Therefore, aspects of the Bandura et al. study that could gain credit here included a definition of aggression and social learning theory. Credit could be given to generic descriptions of the principles for both of these. However, many responses focused too narrowly on the aims of Bandura et al. and what was reported. These could only gain one available example mark.

- (b) The responses to this question varied a lot. Strong responses could clearly show understanding of confidentiality using an example from the study by Bandura et al. Popular explanations included that we only knew the nursery and how many children. For deception, some responses could explain that the children thought the toys were for them or that they did not know they were being observed in the final part of the study. A significant number of responses answered this with an informed consent example, and this could not gain credit as it was not the ethical guideline listed in the question. The guideline of protection from physical harm was known by a significant minority of responses as many tended to be incorrect in their knowledge about this guideline. Many responses stated that nothing happened as the children had not been 'attacked' or that the adults were the only ones performing aggressive acts. Both of these are incorrect as the first would never happen and the second did happen as the children reproduced physically aggressive behaviours. Protection from psychological harm was a stronger guideline in terms of correct responses. The fact that the children were exposed to aggression meant they could have negative long-term effects of being in the study. To improve, candidates need to be explicit with their knowledge of an ethical guideline and then provide an example from the study that shows it was broken or not broken. Each guideline is effectively marked out of two.

Question 10

The strongest responses evaluated the study by Yamamoto 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 laboratory experiments. Common choices included generalisability, quantitative data, and ethics. These strong responses could explain why an element of the study was a strength or a weakness, using specific examples from the study by Yamamoto 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 Yamamoto et al. as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these types of questions as it is testing their evaluation skills only. In addition, some responses followed a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics) and some responses appeared to be prepared essays for Yamamoto et al. without one of their points being explicitly about laboratory experiments. A response that fails to have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses used laboratory experiments 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. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples from the study to show clear understanding. These are the requirements for a Level 4 response.

PSYCHOLOGY

Paper 9990/13 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 the responses. The essay (final question) requires four evaluation points to be present in depth (two strengths and two weaknesses) with at least one of these about the named issue. Credit is limited if the named issue is omitted or just described.

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

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

Candidates need to know about real-world applications for all core studies. To show understanding, answers need to tell the Examiner what the application is based on in the particular core study and then how this could be achieved. Again, this must be explicitly made clear by the candidate.

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

Candidates also need to know the set procedure of studies in the order presented in the original journal article. Questions can be based around just part of a procedure and the candidate must be able to produce an answer that is directed and concise rather than writing about the whole of the procedure. This can sometimes mean a candidate may run out of time for other questions.

There is enough time for answers to be planned to ensure that the response given by a candidate is focused on the demands of each question. This is a crucial skill to develop as some candidates appear to have good knowledge of a study but do not apply this effectively to the question(s) set.

General comments

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

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

There was a significant number of blank responses on this paper. 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 stated the approximate number of participants diagnosed with AS/HFA used in the study by Baron-Cohen et al. Common incorrect responses included a sample size from a different group of participants used in the study, for example, the general population.
- (b) Half of the responses could correctly identify the other judgement. Common incorrect responses included age and emotion, but the latter was in the question. There was a significant number of blank responses for this question.
- (c) There were some correct responses to this question. There was a significant proportion of responses that gave other changes to the eyes test not related to the range of scores being increased. It is important for candidates to carefully read what the question is asking for. There was a significant number of blank responses for this question.

Question 2

- (a) There were not many correct responses to this question. Some candidates could name two objects that a chimpanzee helper could choose from with popular choices being brush and rope. However, a majority of responses were incorrect including stating straw or stick (which was in the question) or other objects not related to the study. Candidates need to carefully read what the question is asking for.
- (b) There were some clear, concise responses to this question. For example, noting the potential issue of generalisability due to using just five mother-child pairings. Other popular choices included a description about the lack of mundane realism, and that the chimpanzees had been used in previous studies. There was a number of responses that did not provide an example from the study to elaborate on a chosen weakness. Candidates need to be aware that 'in this study' in the question requires an explicit example from that study to be able to access all available marks. Also, there was a significant number of responses describing an ethical weakness or arguing that it was cruel to keep chimpanzees in captivity. Neither of these types could not gain credit here as they were not answering the question set.

Question 3

- (a) There were some correct responses to this question. Popular descriptions included the sample size, how many males and female participated and how many were studied intensively. However, there was a significant minority of responses that could not describe three features/characteristics of the sample, or that confused the sample with a different Core Study. It is important for candidates to note the number of marks assigned to a question as this typically represents (in short answer questions) the number of correct elements that need to feature in a response, in this case, three.
- (b) A small minority of responses could describe a result from REM awakenings. These could describe the difference, and which sample the result had come from (e.g., the five participants studied intensively). With results, where possible, there must be a meaningful comparison to be able to access the marks available. Therefore, simply stating that there was less no dream recall in the REM awakenings could only be awarded partial credit. Several incorrect responses chose a different result to describe (for example, the number of dreams recalled overall rather than first half compared to second half) or a result based on the descriptions of dreams given by participants. The important part of this question was that it was about no dream recall, but many candidates wrote about dream recall. It is important for candidates to read the question carefully. There were many blank responses to this question.

Question 4

- (a) Stronger responses could clearly describe four elements of the part of the procedure from the preliminary run until 300 v in the study by Milgram. Popular elements included reading a word and four responses, the increase in voltage per incorrect answer, and what happened when there was no response from Mr. Wallace. Incorrect responses tended to focus on the procedure after the 300 v or the sample shock given to the participant. Some responses simply outlined another part of the procedure in the study by Milgram or the findings.
- (b) For this type of question, responses must contain two parts. The first is the 'what' – what real-world application could be. The second is the 'how' – how will the real-world application be achieved in an ethical way. Common responses included utilising uniforms or some form of verbal prods to encourage the general public to be more obedient. There were some unethical responses about physical punishment, and these were not given any credit.

Question 5

- (a) The majority of responses could outline the aim of the study by Pepperberg. The most popular choice was to test the concept of same/different in an avian species. There were some responses that only gave a very brief aim which could only be awarded partial credit. Some responses presented a finding from the study so could not gain any credit. It is crucial for candidates to read questions carefully.
- (b) Stronger responses could clearly explain one reason for standardisation. The most popular explanation focused on reliability and replicability with an example from the study by Pepperberg. Some responses were awarded partial credit for explaining a reason but not providing the final 'in this study' part of the question. Some responses focused on control and cause and effect and gained appropriate credit. The question was why and not how so responses that simply provided examples of how the study was standardised could not be awarded credit.

Question 6

Stronger responses could clearly outline what is meant by the cognitive approach and provide clear examples from the study by Andrade. Popular descriptions included how we process information, and that cognitive psychology is about how our memory systems work. Popular choices for examples from the study included how we can dual-process information, and that doodling did affect memory recall. To improve, candidates need to have examples from each Core Study that appropriately support each of the approaches at AS-Level. There was a significant number of blank responses to this question.

Question 7

There were some strong responses to this question. Some candidates could clearly describe the procedure between the two time points noted in the question. Popular elements included the doctor leaving the room, a stooge entering the room, and that they had to wait 20 minutes for the Suproxin to take effect. Some responses provided information about a different part of the procedure including the Anger condition or what happened after the stooge started their routine. Candidates need to know the procedures used in all Core Studies. There was a significant number of blank responses to this question.

Question 8

- (a) (i) There were many correct responses to this question. Common choices included 'sock him' and 'Pow'. There was a significant proportion of responses that gave examples of non-imitative verbal aggression or physical aggression. It is important for candidates to carefully read what the question is asking for.
- (ii) There were not many correct responses to this question. The most popular choice was about Bobo being a tough fella. Incorrect responses tended to give an imitative verbal aggression response, or a made-up response not reported by Bandura et al. This question had the highest number of blank responses on the paper.

- (b) There were a few strong responses provided that fully engaged with the stimulus material. Gloria was the most popular choice for the debate by a significant margin. Popular arguments supporting Gloria included the lack of mundane realism with examples such as watching a model 'play' with a Bobo doll. Popular arguments supporting Arturo tended to focus on the controls meaning cause and effect could be established, and that participants were matched on aggressive levels prior to the study. Some responses simply described aspects of validity without engaging in the Gloria-Arturo debate so could only gain partial credit. To improve, candidates need to clearly choose one side of the argument and then explain why they support it using evidence from the study.

Question 9

- (a) For these types of questions, responses should focus on the general psychology that is being investigated in the study rather than a specific aim of the study. Therefore, aspects of the Laney et al. study that could gain credit included a description of false memories and how these can be created, in generic terms. However, many responses focused too narrowly on the aims of Laney et al. and what was reported. These could only gain one available example mark.
- (b) The responses to this question varied a lot. Strong responses could clearly show understanding of confidentiality using an example from the study by Laney et al. One popular explanation was that we only knew they were students at a specific university. For deception, some responses could explain that the participants were not told the truth about their childhood eating habit computer report. A significant number of responses answered this with an informed consent example, and this could not gain credit as this was also an ethical guideline listed in the question. The guideline of informed consent was known by a significant minority of responses as many tended to be incorrect in their knowledge about this guideline. Some responses could differentiate between deception and informed consent by stating that the latter was about giving the participant a choice but that they had not been informed of the true aims of false memories. Protection from psychological harm was a stronger guideline in terms of correct responses. The participants may not have been stressed by simply looking at slides of food, but it could have been argued that the participants might have felt some psychological discomfort from being 'tricked'. To improve, candidates need to be explicit with their knowledge of an ethical guideline and then provide an example from the study that shows it was broken or not broken. Each guideline is effectively marked out of two.

Question 10

The strongest responses evaluated the study by Piliavin 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 quantitative data. Common choices included generalisability, 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 Piliavin 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 Piliavin et al. as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued, with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. In addition, some responses followed a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics) and some responses appeared to be prepared essays for Piliavin et al. without one of their points being explicitly about quantitative data. A response that fails to have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did use quantitative data in an evaluative sense but did not fully explain why it could be a strength and/or a weakness, or simply described quantitative data collected by Piliavin et al. Several responses did not cover the named issue. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples from the study to show clear understanding. These are the requirements for a Level 4 response. There were a significant number of blank essays on this paper.

PSYCHOLOGY

Paper 9990/21 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.

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.

General comments

In general, candidate responses achieved marks across the whole range of available marks for this paper. However, very few responses consistently and accurately demonstrated knowledge and understanding or achieved the additional marks for linking the response to the scenarios, thus limiting marks achieved overall. Nevertheless, some of the candidate responses showed a good grasp of a range of psychological concepts and, therefore, achieved the basic marks.

Nearly all of the questions required a link to a study. These links are of different types:

- Some links are to a specific key study. These included **Questions 1(a), 1(b), 3(b)**.
- Other links are to the information provided in the stem (the introduction) to the question. These included **Questions 2, 4(a)(i), 4(a)(ii), 7(a), 7(b)(i), 7(b)(ii), 8(a), 8(b)(i), 9(a) and 9(b)**.
- A link or use of any examples from any study also appears. This type included question 6.

If a question required a link, of any type, and that link is absent, then only partial marks can be awarded. This often results in 1 mark being awarded rather than 2, and sometimes 2 marks rather than 4.

There were not many strong responses to **Question 10**. Candidates could often explain a natural experiment, the difficulty was in applying it. Many candidates manipulated an IV, which is incorrect. Candidates did not consider a public event, as the question required. Candidates should be able to plan a study for all methods that are on the syllabus, and it is very important to read the specific requirements of the question.

Comments on specific questions

Section A

Question 1

- (a) In the study by Bandura et al. quantitative data was collected using observation. Inclusion of the word 'observation' scored 1 mark and if this was linked to the study by Bandura et al., a second mark could be awarded. For example, if a candidate wrote 'observation (1 mark) of the number of times a child copied the model' (2 marks). Some candidates defined quantitative data, 'the number of times a behaviour was recorded', but this could only be awarded a partial 1 mark because it does not describe how quantitative data was collected.
- (b) Nearly all candidates scored 1 mark for suggesting an advantage of quantitative data such as 'it can be analysed mathematically/statistically'. However, many candidates scored no further marks because their answer was not linked to the study by Bandura et al. as the question required. Many candidates did score full marks when stating for example that 'they could compare the mean aggression of boys and girls' or when stating 'they could calculate the percentage of children copying aggressive models'.

Question 2

- (a) Although many candidates were awarded full marks, others scored no marks at all. Answers achieving no marks were for two reasons: first, the stem stated 'experimental hypothesis' and so any correlation was incorrect. Second, those defining directional hypothesis as 'stating the direction' with no elaboration beyond this were merely re-writing the question. Correct answers were 'Because it says which way the difference will go' (1 mark) and for the second mark 'that older people will forget more than younger people' (2 marks). It was also possible to define directional hypothesis by explaining how it differed from a non-directional hypothesis.

Question 3

- (a) The majority of candidates scored 1 mark for correctly describing an uncontrolled variable. Answers such as '(problem) variables that have not been, or cannot be successfully eliminated,' were common. A few candidates merely repeated the question 'an uncontrolled variable is a variable that is uncontrolled' and such answers scored 0 marks. A few candidates gave an example and whilst this confirmed their understanding, an example was not required as this was a 1-mark question.
- (b) This question caused problems for a number of candidates who confused doodling with daydreaming. Other candidates could state the problem but often did not go on to say how this would affect the result. For those being awarded 2 marks, daydreaming was explained as an uncontrolled variable (i.e. which could not be controlled) and that the effect was that confounding was likely (i.e. unsure whether the DV was caused by the IV or some other variable).

Question 4

- (a) Some candidates defined a positive correlation for (a)(i) and some defined a negative correlation for (a)(ii) both of which scored zero marks. Both question parts had to be related to the stem of the question which related the amount of time spent using a phone and the quality of sleep. A perfect answer for (a)(i) would be 'when people spend more [less] time using their phone they have a better [worse] quality sleep'. For (a)(ii) a perfect answer would be 'people who spend more [less] time using the phone have a poorer [better] quality sleep'.
- (b) There were a number of possible answers to this question. The first might be that an experiment would be very difficult to conduct because participants might not wish to change their phone behaviour for example, or because there would be too many variables that could not be controlled. A second could be that sometimes correlations are conducted before an experiment. A third might be that the study is looking for a relationship (between the two variables of phone usage and quality of sleep) and not causation.

Question 5

- (a) Most candidates scored 0 marks in response to this question mainly because the specific question asked was not answered. For example, the question focused on comments and not helping behaviour (as many candidates assumed). Second, the question was not about the comments made by the participants, what was actually said or whether these were positive or negative. The question focused specifically on the differences between the type of victim. Answers such as 'there was a difference in comments made in relation to black and white victims' scored 1 mark; answers such as 'more comments were made when the victim was white than when they were black' scored 2 marks. Similarly, this type of answer applied to 'drunk' and 'cane' type victims.
- (b) Nearly all candidates scored 1 mark by providing an appropriate problem experienced by the observers. For example, the observers might have miss heard what was said; might not have understood, if participants were speaking a different language; might have misremembered what was said by participants.

Question 6

This question required a description of a semi-structured interview and an unstructured interview. Answers were of four types:

- correct answers stating that semi structured interviews follow some set questions, but the interviewer can add some questions in response to the participant's responses; that unstructured interviews have no set questions.
- answers which assumed that semi structured interviews automatically provide quantitative data and unstructured interviews provide qualitative data. The type of interview has nothing to do with the type of data that is gathered.
- answers where structured interviews were confused with semi-structured.
- those where no examples were provided, or where examples were far too vague rather than being explained (for example, stating nothing more than 'for example, Milgram' or 'a job interview').

Question 7

- (a) Most answers were awarded 1 mark for correctly stating the aim of the study which was to investigate (or test) how much time different size babies spend in dream sleep. A small number of candidates stated, 'to investigate how much time babies spend in REM sleep', ignoring the size of the baby bigger or smaller altogether, and so scored 0 marks.
- (b) (i) The majority of candidates scored full marks in response to this question. Typically, it was stated that having electrodes on the baby's head would not be comfortable for them and their sleep would be disrupted.
- (ii) Any technique that was not uncomfortable or would disrupt sleep or cause psychological harm to the baby was creditworthy. Simply observing eye movements would be appropriate. Some candidates suggested using an EOG as an alternative to EEG, but as an EOG also involves electrodes this would not be a more appropriate technique. Some candidates suggested using an fMRI, but this would also be inappropriate because babies would not remain still whilst being scanned.
- (c) This question part focused on the content of dreams unlike the above question parts which focused on REM sleep duration. This meant that answers suggesting observing eye movements needed to include the direction (vertical or horizontal) of the movements suggesting the content of the dream. Also creditworthy was any noises or facial expression that a baby might make during REM sleep which might indicate dream content. The observation of body movements scored 0 marks because the body moves during NREM sleep and is relatively paralysed during REM sleep.

Question 8

- (a) The question asked for a reason 'why the elderly person has an animal' and asked for 'one other reason for choosing the participant'. Many correct answers stated simply 'a dog because it is a good companion for an elderly person' and were awarded 2 marks. Some candidates did not name any animal at all, and some identified an animal but could not give a reason why an elderly person would have an animal. Regarding 'the other reason' some candidates stated that the participant needed to be elderly, but this scored no marks because that was known from the stem. Defining an age or age range was creditworthy, as was suggesting that the participants should live alone, and linked this to the dog being a companion to avoid loneliness.
- (b) Question parts (i), (ii) and (iii) were linked. **Part (i)** required the identification of a technique (such as a questionnaire), part (ii) required an advantage of this technique and part (iii) required a disadvantage. Most candidates were able to answer all three question parts, but often could only be awarded partial marks because the answer was not linked to the particular case study as the question required.

Question 9

- (a) Many answers achieved the full mark for providing a problem Rory could have when observing fish. Some candidates suggested that the fish might move too quickly for Rory, others suggested that the fish might change their behaviour because of the presence of Rory. Yet others suggested that the fish might hide if there was an object in the tank.
- (b) This question focused on the 'animal ethics' guideline of housing. Most candidates scored 1 mark for an appropriate suggestion such as the water freshness, temperature, or cleanliness. Others focused on the environment suggesting that there should be plants or some rocks, and others mentioned that the tank should be of an appropriate size. Not all candidates scored the second mark, but those that did suggested that including a feature meant that the environment of the fish was more natural or would reduce stress levels.
- (c) Many answers correctly stated that generalisability would be limited because Rory was only studying one type of fish and so any results would not apply to types of fish that were not included in the sample. A number of candidates went further than this, stating that fish live in water and so the findings would not apply to any animal that does not live in water.

Question 10

- (a) The required method was a natural experiment, a method where the IV occurs naturally rather than being manipulated by an experimenter. Many designs involved stooges manipulating the behaviour of participants and such answers scored no marks. A natural experiment can have independent variables. For example, in this instance the IV could be before and after a positive public event (with the event itself not manipulated). Similarly, candidates could have compared a positive public event with a negative public event. Another weakness in answers was that the event chosen was not a public event. Many candidates chose an event involving school friends or a family event, who are people who know each other, and are not members of the public attending a public event. Examples of types of public events were included in the stem of the question, which were often ignored in favour of an incorrect alternative. Another weakness in answers was that the DV of friendliness was often vague with 'I would measure how friendly people are' without any clarification of how friendliness would actually be measured. Some candidates even decided to measure aggression which was irrelevant. Candidates should always answer the question that is set and planning their answer before beginning to write is advisable.
- (b) Despite the question stating: 'do not refer to ethics or sampling', a number of answers incorrectly focused on one of these aspects. Further, the question required a limitation of the procedure, and many answers did not consider a limitation of the procedure. There were also many excellent answers. For example, some answers mentioned the use of inadequate or vague response categories, meaning that some aspects of friendliness could not be recorded when behaviour was observed. A frequently suggested solution was to conduct a pilot study to determine adequate categories.

PSYCHOLOGY

Paper 9990/22 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 abilities, although application of knowledge to unfamiliar contexts in short answer questions presented a challenge for many candidates. This is an essential skill and candidates benefit from the opportunity to learn to respond to scenario-based questions. This is particularly in relation to two skills; extracting information from scenarios and linking responses directly to the content of the scenario described.
- Grounding in basic concepts is critical for success in research methods. Responses to this paper suggested that some candidates did not fully understand all concepts. Examples of this included:
 - the differences between different types of questions, including providing the 'anchors' for rating scales and using interrogatives that lead to detailed answers in open questions (such as 'describe' or 'explain why')
 - the similarities and differences between different types of interviews
 - the similarities and differences between types of experiments.
- In addition, candidates should read and follow the instructions given in the question carefully. It is important to read the whole of the stem and the question, and ensure that they are answering the question asked.

General comments

Candidates were able to demonstrate knowledge of a range of aspects of research methods in this paper. Success was greater on straightforward and low mark tariff questions such as **1** (providing a basic definition of reliability), **2** (making a simple distinction between qualitative and quantitative data), **3(b)(ii)** (explaining an ethical issue in a somewhat familiar context) and **8(a)(i)** (writing a scale for 1 mark). Candidates were also successful on **Question 8(b)(i)** which was based on a novel scenario. This question required detail and led unavoidably to providing a linked answer. In this instance, candidates were able to apply their knowledge successfully. Most candidates were also able to score some marks on more complex questions, but less often earn full credit. In questions such as **3(b)(i)** and **4(b)** elaboration of the answer was required for full marks, and in **Questions 7(b)(i)** and **9(d)** the response needed to be linked to the scenario to gain full marks.

One question which presented candidates with a considerable challenge was **Question 7(a)**, responses to which indicated a poor understanding of the axes of a normal distribution curve.

Question 10 had some strong responses, although many responses suggested that the candidate was unclear about how a natural experiment differed from a field or laboratory experiment. As a result, responses either omitted important elements (e.g. the independent variable) or included elements that should not have been present (e.g. stooges who manipulated the experimental situation).

There were almost no issues with legibility and use of additional paper was almost always clearly signposted.

Section A

Question 1

This question was generally well answered with most candidates achieving one mark. Common errors included simply stating that reliability is about consistency, referring to the reliability of results without mentioning the procedure and, in a small minority of cases, confusing reliability with validity or referring only to 'accuracy'. It is important for candidates to understand that reliability is not just about the findings, rather the procedure, research, measures, etc.

Question 2

This question was very well answered, with many candidates achieving full marks. A minority of candidates confused qualitative and quantitative data. Other common mistakes were typically stating that qualitative data is subjective / opinions / about feelings / in words and that quantitative data is objective.

Question 3

- (a) There was evidence that candidates found this question challenging, with very few achieving the full two marks. Common errors included focusing on the behaviour of the stooge, writing generic responses which were merely a repetition of the stem and failing to link clearly to the 'effect'. Other frequent mistakes were to discuss only the gender of the stooge rather than that of the participants, or to give a generic answer that males/females would behave 'differently', thus simply repeating the stem.
- (b) (i) Many candidates achieved at least one mark for correctly identifying a participant variable but were unable to contextualise this in terms of the effect on the results. Quite a few responses did not answer the question, as they did not identify a participant variable. A common mistake was to ignore the part of the question asking for a participant variable, and to instead just respond with what they thought the results of the experiment would be more generally – e.g. that the ill stooge would make participants feel ill.
- (ii) Most candidates answered this question part well. They could identify a suitable ethical problem and the majority could link that clearly to the second experiment, with the most common issue identified being deception.

Question 4

- (a) This question part was generally answered well. However, some candidates did not earn credit on this part of question because they repeated the question, stating that there were only two options. This sometimes led to them being unable to express themselves in the way they wanted to in the second part of the question. The term 'ceiling effects' was often simply repeated in both parts of the answer without any further explanation, meaning candidates could not access the full four marks. Some candidates were unable to demonstrate sufficient understanding of the study and stated that one of the eight problems identified by Baron-Cohen et al. relating to the forced choice question was that the 'options did not offer the choice that the participant wanted to make'.
- (b) Generally candidates achieved at least one mark here, for correctly identifying an increase in response options. However, this was not always followed up with a clear explanation of how it improved validity.

Question 5

The most common mistakes were to give generic answers rather than specific occupations, such as 'professionals' or 'blue/white collar workers' or to give characteristics of the sample (male, age, from New Haven, etc.). Where occupations were given, they were typically incorrect, for example a doctor, accountant or businessman.

Section B

Question 6

This question was generally well answered with candidates accessing the full range of top marks. However, it is important for the candidates to understand that closed questions must be written with their answer options and any scales must be defined (e.g. giving the 'anchors' such as 0 = not at all, 10 = completely). In addition, open questions should clearly elicit detailed answers, for example by including the words 'describe', 'explain', 'why', etc.

Question 7

- (a) Most candidates scored some marks on this question part, however few achieved full marks. Many candidates appeared not to know what normal distribution should look like and included for example line graphs, scatter graphs and bar charts in their response. The biggest loss of marks came from incorrectly labelled axes with very few participants understanding that the y-axis needed to refer to frequency or 'the number of participants'. Most candidates incorrectly labelled the y-axis with 'positivity levels'.
- (b) (i) Most candidates could identify a practical limitation of structured interviews, however there was a significant minority who suggested that structured interviews consist only of closed questions, so gave incorrect responses such as 'participants can't explain their answer'.
- (ii) Many candidates correctly identified invasion of privacy with a few referring incorrectly to confidentiality. Where candidates did not score full marks, this was typically because they did not link their answer to the context, i.e. did not explain why or how the issue of invasion of privacy might be relevant to research on social media use.

Question 8

- (a) (i) This question part was typically well answered. However, where the mark was not earned, the most common issues included not giving the 'anchors' of the scale, suggesting a Likert scale with no statement, or just suggesting 'doing a rating scale' without a suggestion of what that scale might be. A very small minority of candidates simply stated 'use a feelings thermometer', without any detail.
- (ii) This question part did not have many strong responses, with candidates often focussing their response on the teachers (or the general procedure of the study) rather than the scale. More commonly, however, they suggested that the scale does not tell us 'why'. Note that subjectivity is an inherent feature of some measures, which is reflected in the participants' choices on the scale. Many other answers were generic, i.e. unrelated to the specific scale suggested by the candidate in the previous part. In addition to the mark scheme examples, candidates raised relevant points such as the absence of a midpoint option on their scale meaning that the teachers could not give a 'neutral' level of annoyance.
- (b) (i) This question part was very well answered. There were many inventive (and good) responses, indicating that students were effectively applying their knowledge of self-reports to the scenario given. Most commonly students referred to social desirability bias and demand characteristics. It was clear that a large number of students were using these terms interchangeably and so perhaps do not understand the difference between them.
- (ii) Where good responses were given to this question, many were similar to those on the mark scheme. However, an alternative good answer that was commonly suggested was that the knowledge of the difference may have distressed the teachers as they would doubt their own level of self-awareness. Where the mark was not earned, this was typically because the candidate had just named a different aspect of harm – guilt/embarrassment/hurt without explaining 'why'.

Question 9

- (a) Most candidates identified species (and strain) and housing with many of them able to elaborate in the context of the study. Some candidates referred incorrectly to 'numbers', 'replacement' and 'protection from harm'.

- (b) There was evidence that most candidates knew how to deal with two middle numbers and most earned the mark. Where this was not the case, candidates often gave the answer to **9(c)** (they described how to find the range) or suggested 'finding the average' of the two middle numbers. Since the 'average' could refer to the mean, the mode, or to the median itself, this was not a sufficiently clear answer to earn credit.
- (c) Identifying the range appeared more challenging for many candidates. However, many scored one mark, often missing a link to sociable behaviours from their answer and merely giving a generic response about how to calculate range.
- (d) Candidates appear to know that counterbalancing reduces order effects, such as practice effects and fatigue effects. However, the majority were unable to apply this to the scenario, often merely explaining how counterbalancing would be carried out without referring to why it would be necessary to do so in this instance. In other words, they were answering the question 'how would counterbalancing be used' rather than the question asked, why it was appropriate. Nevertheless, some candidates were able to earn full marks, for example by saying that it was necessary to counteract order effects because the animals might have become more familiar with each other by their exposure to the second condition, so would be more sociable.

Section C

Question 10

- (a) There were many good examples of natural experiments using observation or questionnaires in response to this question part. Stronger responses gave clear examples of behavioural categories and examples of questions. Most opted for the trolley/basket option as detailed in the stem, but some opted for more creative differences which worked equally well. Some candidates were unclear about natural experiments, being unable to identify the IV/DV or inappropriately described stooges who engaged with or tried to annoy the participants in order to elicit behaviour which could then be recorded. Others simply referred to observers as 'stooges', even though they were not. These aspects of the planned studies were ignored as they did not follow the rubric of describing a natural experiment. Another common misconception was that there is no independent variable in a natural experiment, rather than the IV not being manipulated. In such circumstances, candidates would only state the DV, but in spite of this, many were able to compare two conditions.

Further weaknesses in the candidates' responses included:

- A problem with teenage vernacular means the use of the word 'random' is often inappropriate. It is important that the candidates understand this difference.
- Stating an aim/hypothesis which may have helped the candidate to focus on the question, but in itself would not earn credit.
- Many candidates used 'time spent in the shop' as a measure of the DV. Often this was without an explanation of how this linked to relaxation. Stronger answers operationalised a lack of relaxation as 'rushing' (short time) or as 'dithering' (long time).
- In producing behavioural categories for observation of whether shoppers were relaxed, candidates encountered two problems. Firstly, they often suggested emotions and since states cannot be observed, these were not creditworthy. Stronger responses suggested observing indicators or relation of tension, such as smiling and frowning. In addition, there was a tendency to use measures that were not – or not necessarily – indicators of relaxation, such as 'time spent shopping' or 'number of items bought'. Stronger responses operationalised relaxation by suggesting that 'relaxed shoppers moved slowly'/'tense shoppers scurried' so fast speed was associated with a lack of relaxation. Similarly, justifications were given for number of items in stronger responses.
- One additional, and moderately frequent, mistake made by candidates in answering this question was to include 'data'. This included, for example, explicitly quoting the number or percentage of 'trolley' and 'basket' shoppers who were relaxed which the study 'found'. Sometimes this constituted the majority of a candidate's response. Such information is not answering the question, so does not earn credit.

The majority of responses fell into Level 2. Few achieved the very top marks of 9 or 10/10.

- (b) There were many good answers to this question although a minority of candidates did not follow the rubric about sampling / ethics trap so could not earn credit. Another common error was that candidates had described a field experiment in their procedure and their solution was to make their procedure a natural experiment. There were many good generic answers which could have been stronger answers if they had been applied to the candidate's answer to **part 10(a)**.

A significant error related to reliability. It is meaningless to talk about inter-observer reliability when there is only one observer. Additional researchers do not, per se, improve reliability. Inter-observer reliability can only be improved when two or more observers are compared and take action to make their observations more similar.

PSYCHOLOGY

Paper 9990/23 Research Methods

Key messages

- Paper 2 is a research methods paper asking candidates a range of questions. Some of the topics include the core studies in relation to aims and hypotheses, sampling, independent and dependent variables, control conditions, data analysis and ethical guidelines in relation to humans and animals. In addition, more general questions are asked about research methods, terms and concepts used to describe or evaluate research methodology, and application of this knowledge to both familiar and unfamiliar contexts. Such questions require different skills, and therefore it is essential that candidates are prepared for each of these skills, especially recalling concepts and the application of this knowledge.
- Applying knowledge and understanding to novel scenarios is important to succeed on this paper. This can help candidates in two ways:
 - Candidates need to be able to apply research methods, terms and concepts to scenarios presented in questions. These can include, for example, planning, criticising or developing designs or analysing data.
 - Candidates must take note of questions which indicate the need for a link. When a question says 'in this study', or makes direct reference to the scenario, responses must go beyond simply describing or evaluating, they must contextualise the answer in a relevant way. Candidates therefore need to be prepared for questions using this format and practice can help them to learn both how to extract relevant ideas and how to make novel suggestions based on scenarios.
- It is important that candidates have a good grounding in the basic principles of research methodology, for example, the different research methods, sampling methods and techniques of data analysis. Without these, candidates will find it difficult to answer even basic questions and application questions will present an even greater challenge.

General comments

Candidates were most successful on questions that drew on understanding from other subject areas – such as knowledge from mathematics to aid graph drawing in **Question 8(a)(i)**, although most were not able to access more specific areas of such questions. Candidates performed well on questions where they could utilise more general subject knowledge from psychology, such as in **Question 7** relating to learning, rather than relying exclusively on their knowledge of research methods. Very few candidates were able to accurately or consistently demonstrate knowledge and understanding, or to access additional marks for linking their response to the scenarios. This limited their performance.

Candidates were able to demonstrate some knowledge of research methods in this paper. Success was greater on more straightforward questions such as **Questions 2(b)**, changing a non-directional hypothesis into a directional hypothesis; **3(b)**, identifying the dependent variable in a familiar study; **7(a)(i)** and **7(b)(i)**, as mentioned above; and **Question 8(a)(i)**, as mentioned above.

Question 10 was sometimes well answered, although many responses were exclusively or partly based on methods other than the required interview, such as questionnaires. This meant that the response lacked the necessary relevant detail to earn credit as the candidate ran out of time.

Comments on specific questions

Section A

Question 1

There were many irrelevant responses to this question, often referring to general aspects of the procedure rather than those which specifically helped to reduce demand characteristics, or to other aspects of the study unrelated to the question, such as the aims.

Question 2

- (a) There were very few correct answers to this question. Typically, candidates described the aim and hypothesis in the same way but in different words. One common incorrect statement defined a hypothesis as 'an educated guess'.
- (b) This question was generally well answered, and the majority of appropriate responses earned credit as they referred to both boys and girls. A small number of attempted hypotheses were incomplete, for example saying only 'Boys will be more aggressive'. However, there was also a range of incorrect responses including explanations of Bandura's study, descriptions of various irrelevant differences between boys and girls, and other unrelated comments. Where a hypothesis was offered, it was sometimes a null hypothesis or a rewriting of a non-directional hypothesis.

Question 3

- (a) There were few correct responses to this question part. Incorrect responses typically included describing the DV or just stating 'office'.
- (b) This question part was answered better than **part 3(a)**, with responses either in terms of obedience or its measurement.

Question 4

- (a) There was a range of responses to this question, from fully successful to unsuccessful. '*Don't use endangered animals*' was a common one-mark answer, but it was rarely supported by explanation. A frequent incorrect response was to suggest that animals cannot give consent. It is also worth noting that in this context the word 'strain' does not refer to 'stress' but to a variety of species which is generically similar. Many candidates were unaware of the meaning of this term in the syllabus.
- (b) Some answers to this question part gained 1 or 2 of the three marks available, most commonly for ideas relating to 'food', 'space' and 'safety'. There were also many answers that suggested the candidates had insufficient knowledge of this guideline. A frequent mistake was to suggest that rather than the actual purpose of the housing guideline – to create a safe and healthy environment – it aimed to try to recreate the wild or natural environment.

Question 5

- (a) Some candidates gained the mark, typically for saying 'it's where the IV is missing'. Responses were often poorly expressed, even those that were creditable.
- (b) Where candidates were able to gain marks here, they typically earned at least one of the two marks for identifying the control condition in Schachter and Singer's study.

Question 6

This question did not receive many strong responses. In particular, random sampling was weakly described by the majority of candidates. Many responses described random sampling as '*finding random participants*' and volunteer sampling as '*asking people to volunteer*'.

Section B

Question 7

- (a) (i) This question part was often well answered, with many candidates gaining full marks. Common answers were 'homework passes', 'time off' and 'pizza'.
- A small number of candidates only described when teachers should give rewards without including what those rewards could be.
- (ii) This question part was also well answered, with many candidates gaining at least one mark. Some excellent responses related to the candidates' ideas were given, for example: 'homework passes'/'time off': less work done than alternative so hinders learning rather than helping it; 'pizza': some students might not like it, so less effective than alternative suggestion.
- (b) (i) This question part was often well answered, with most candidates gaining at least one mark, typically for identifying a relevant guideline. Many were also able to earn the second mark for giving detail, typically relating their answer back to the scenario.
- (ii) This question part was less well answered than **part (b)(i)**, but nevertheless, many candidates gained at least one mark for identification of an appropriate practical problem. A small number of candidates referred again to ethical problems so could not earn credit.

Question 8

- (a) (i) This question part was the best answered on the paper, with many candidates gaining full marks.
- (ii) This question part, for which the answer was 'mode', was poorly answered. Incorrect answers were varied, including foods, colours, percentage, word type and bar chart. Only a minority of incorrect responses were the mean or median. This suggested that many candidates were unfamiliar with the term 'measure of central tendency'.
- (b) There were very few correct answers to this question part. Candidates often supplied a term, such as 'generalisability' without any detail. This was not creditworthy as the question required an 'explanation'.
- (c) There were many good responses here, with many candidates offering at least one correct disadvantage, suggesting that they were able to apply their understanding of conducting observations to novel scenarios.

Question 9

- (a) This question did not receive many strong responses. Many responses did not differentiate natural experiments from field experiments. Most answers related to the participants 'natural environment'. As with **Question 8(a)(ii)**, many candidates appeared unable to demonstrate sufficient psychological knowledge. Where candidates demonstrated a somewhat better understanding, they were often unable to correctly identify what was 'not manipulated', e.g. they tended to say '*nothing was manipulated*' or '*it was not manipulated*'. Finally, even those candidates who could define a natural experiment were often unable to provide a relevant example.
- (b) This question was well answered, with candidates typically identifying a generic problem. Fewer candidates were able to link this to the study, so could only gain one of the two marks available.
- (c) (i) Many candidates were typically able to gain one mark on this question part. However, the majority of candidates did not earn full marks, largely because – as with **part (b)** – they did not link their answer to Dr Felix's study, i.e. did not respond to the end of the instruction '...in this study.'
- (ii) Candidates were not as successful on this question part as with **part (c)(i)**, often giving irrelevant responses.

Section C

Question 10

- (a) This question part did not receive many strong responses. Many candidates did not describe interviews. The most common incorrect answer was to describe questionnaires, but candidates also offered lab experiments, field experiments and case studies. When interviews (or questionnaires) were described, the suggestions for what would be asked were often written as '*I would ask about colour/meaning/consistency in their dreams*', i.e. repeating the aim, rather than suggesting an actual wording for a question. Those questions that were given were almost always appropriate, however, they were rarely identified by format (open or closed) and when they were they rarely matched the open/closed label they were given. In addition, the interview format was rarely identified (structured / semi-structured / unstructured), although whether the interview used a face-to-face, telephone or online technique was occasionally stated.

The candidates tended to use word 'debrief' when they meant 'brief'.

- (b) Few candidates gained full marks on this question part. Commonly, this was either because they did not make their answers specific to the procedure they had described in **part (a)**, or because they offered only a limitation rather than a limitation and a way to overcome it.

PSYCHOLOGY

Paper 9990/31 Specialist Options: Theory

Key messages

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

It is important that candidates are aware of the terminology, theories, and disorders identified in the syllabus as some were unable to identify and/or define the terms/concepts given in these types 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.

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

These questions could ask the candidate to describe a theory, study or treatment(s) used by psychologists that 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 and a result, 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, Bipolar and related disorders: treatment and management of depression: rational emotive behaviour therapy (Ellis, 1962) to help the candidate identify which part of the syllabus the question is referring to.

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. This question is worth 6 marks so the candidate should write a more extended answer for each issue raised. Some responses were very detailed for one issue but then only briefly discussed the second issue. In addition, many of the responses were general and not specific to the study, theory or technique(s) named in the question. To improve, responses should give specific examples to support their point. As mentioned for the odd question **part (b)**, the candidate should make a flashcard / revision notes and could include in this strengths and weaknesses of the theory, study or technique to help candidates prepare for these questions.

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

This question will always come from one of the bullet points in the syllabus. Candidates could describe the three (or four) studies, theories, explanations of disorders or techniques identified in the specification under the appropriate bullet point. For this exam, some of the answers used the incorrect topic area in the syllabus or the description was brief. It is possible for the responses to achieve full marks by describing at least two of the studies, theories, explanations of disorders or techniques and this would need to be a very detailed description. It is also important that the descriptions are linked to the topic area named in the syllabus. It could be useful for candidates to create revision notes with the title of each bullet point as the header in their notes.

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

This question will always ask the candidate to evaluate the theories, studies, explanations of disorders and/or techniques described in **part (a)** of the question. The response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. However, most responses that evaluated using two issues in this exam, achieved in the lower bands due to the response being superficial and often with little analysis. Some responses that considered three issues tended to achieve higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the theories, studies, disorders and techniques described in the **part (a)** of the answer. 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 to show excellent understanding of the issue under discussion. In order to achieve the requirements of the level 3 and 4 band descriptors, it would be best if the response was structured by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure that the answer covers this requirement of the question. Some of the responses covered other issues rather than the one named in the question. Quite a few of the answers were structured by study/theory/technique rather than by the issue, which often led the response to be quite superficial and repetitive. Some of the responses did analysis. Candidates should be aware this question is worth 10 marks and attempt to include an appropriate amount of information.

General comments

The marks achieved by candidates were across the full range of the mark band. Many candidates were well prepared for the exam and showed good knowledge, understanding and evaluation throughout their responses. Some candidates showed limited knowledge and understanding with brief and/or superficial responses. In these responses, there was often evidence of limited evaluation skills.

Time management for this paper was good for the majority candidates and most attempted all questions that were required. Some 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 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.

Comments on specific questions

Psychology and Abnormality

Question 1

- (a) There were many good responses to this question which identified a specific symptom of mania (e.g. increased energy) and a specific symptom of depression (e.g. withdrawal from social activities). Most responses were able to gain one mark for identifying that there was a high and a low mood phase or stating that bipolar disorder involves mania and depression. Some responses identified characteristics of other disorders such as delusions and hallucinations which were not creditworthy.
- (b) Strong responses to this question gave a clear description of REBT with reference to the ABC model (sometimes reference to ABCDE) and how this is used as a treatment for depression. Many responses did not link their description of the treatment to improving the symptoms of depression which limited the marks received. Weaker responses tended to just identify that the therapist challenges the irrational thoughts of a patient.
- (c) The responses to this question covered the full range of the mark scheme. Stronger responses identified a similarity and a difference between REBT and a biological treatment with clear examples from both to explain the comparison point. Common similarities included that both are effective for improving the symptoms of depression, both require a professional and both take time to be effective. Common differences include side effects and the approach taken to treat depression. Weaker responses often identified the similarity or difference without giving any examples to explain these. Common errors included stating that REBT deals with the cause of

depression or that SSRIs take effect immediately which were not creditworthy. In addition, a small number of responses compared REBT with CBT which was also not creditworthy.

Question 2

- (a) Responses varied considerably for this question and covered the full range of the marks available. Some responses highlighted how well prepared some of the candidates were for this exam whereas others showed limited knowledge of this topic. Some responses were detailed, accurate and coherent with a good use of psychological terminology. The strongest responses covered three of the four explanations in detail with some very good descriptions of the behavioural explanation (with the example of the Watson study given) as well as the biomedical/genetic (with the example of the Ost study). Weaker responses often gave more limited descriptions that lacked reference to terminology or clear details of the explanations and/or studies. It was common for the psychodynamic explanation to cover the whole of the Oedipus complex/description of the little Hans' study without much focus on the explanation of the development of a phobia.
- (b) Many of the responses achieved in the level 1 or level 2 mark band, with a few providing clear analysis and details of the explanations of phobias to back up their evaluative points that enabled these type of responses to achieve level 3 and above. There was a tendency for responses to focus on many issues per explanation/study rather than applying the issue to the different explanations/studies. A few responses did effectively discuss the named issue of nature versus nurture with some clear analysis. Common evaluation issues included determinism versus free-will, reductionism versus holism and application to everyday life. Weaker responses tended to identify which side of a debate the explanation was on (e.g. genetic is nature) without any explanation given for why this is the case.

Psychology and Consumer Behaviour

Question 3

- (a) Many of the responses achieved full marks by identifying two of the Cs from Lauterborn's 4 Cs marketing mix model of advertising. Responses did cover all four of the Cs including cost, convenience, communication and customer wants and/or needs. A common error was to state 'customer' on its own which was not creditworthy.
- (b) There were some clear and detailed responses describing the procedure and one result from the study by Auty and Lewis (2004) on product placement in films. There were many good descriptions of the two groups used in the study and how the procedure varied for these participants. Some responses gave details of the prompts used, and also a specific result for the study. Weaker responses often identified that the children watched the 'Home Alone' film and sometimes stated that the children showed a preference for Pepsi when exposed to this during the film but knew few specific details of either the procedure or a result.
- (c) The marks for the responses to this question covered the full range of the mark scheme. Common strengths included generalisability due to two age groups of participants, strengths of quantitative data collected, practical applications of product placement in films and how the study met ethical guidelines. Common weaknesses included generalisability and weaknesses of quantitative data. Stronger responses identified the strength and weakness and gave a clear example from the study. Weaker responses often just identified the strength and/or weakness with either no example given or a very brief explanation. Some responses incorrectly identified a weakness of ethics for the study which was not creditworthy as informed consent was given by the parent/guardian.

Question 4

- (a) The marks for this question covered the full range of the mark scheme. There was a number of good and detailed descriptions of the studies by North et al., Guéguen et al., and Woods et al. Weaker responses often gave fewer details of the studies with a general outline and conclusion given. Some responses gave muddled descriptions for the Woods et al. study.

- (b) Those responses that achieved in the higher mark bands for **part (a)**, produced strong answers to this question with a good outline of the strengths and weaknesses of the experimental method as used in the studies. Generalisability, ecological validity and applications to everyday life were other commonly chosen issues. Weaker responses often attempted a large number of evaluation issues without discussing any of them in depth or giving any examples from the **part (a)** of the response to back up their points. These type of responses often achieved level 1.

Psychology and Health

Question 5

- (a) The vast majority of responses to this question achieved one mark by identifying a characteristic of Type A personality, such as being competitive or controlling. Most responses did not explain why this characteristic could lead to a stress-related illness and often simply stated that being Type A led to experiencing more stress.
- (b) There were many strong responses to this question with many giving a detailed description of stress inoculation training (SIT) as a technique for preventing stress developed by Meichenbaum. Strong responses briefly outlined the three phases of the therapy (conceptualisation, skills acquisition and rehearsal and application and follow through) and linked this to reducing stress. Weaker responses often gave a vague description of therapy that could reduce stress that identified or indicated one of phases or identified relaxation to achieve one mark.
- (c) Most responses identified both a strength and weakness of SIT. The most common strength was effectiveness/developing the skills needed to reduce stress in the future when therapy has finished. Common weaknesses included time/cost and the level of commitment/motivation required from the client. Stronger responses identified the strength and weakness and gave a clear explanation with examples to explain their point. Weaker responses often just identified the strength/weakness without any explanation or example. Some responses identified incorrect evaluation issues such as ecological validity and self-reports which were not creditworthy.

Question 6

- (a) The responses to this question covered the full range of the mark scheme. Stronger responses showed a good understanding of biochemical, psychological and alternative techniques to manage and treat pain. These responses were able to give an outline of how the treatment worked and why it led to a reduction in pain. Many referred to the gate control theory to explain how pain is reduced. Weaker responses often identified the techniques, frequently just identifying medication without naming anything specific, and no indication was given of how these treatments would manage and/or control pain. Stronger responses were given for the psychological techniques where candidates often had more detailed knowledge of these strategies.
- (b) There were some good responses to this question. These responses were often able to evaluate the named issue of determinism versus free-will, and were able to use the treatments from **part (a)** of their answer to explain how each treatment either led to a deterministic reduction in pain and/or the patient enacted their free-will to manage their pain. Stronger responses were able to argue that all of the treatments partly involve the free-will of the patient and why. Other common evaluation issues included cost, time and side effects of medication. Weaker responses sometimes covered the named issue but often just identified whether the treatment was deterministic or down to the free-will of the patient with no example or explanation to back up the point made. Some responses outlined a number of issues but in a superficial manner, simply stating that issues did or did not apply to treatments but not why. These types of responses often achieved in the level 1 mark band.

Psychology and Organisations

Question 7

- (a) There were some good responses to this question with some achieving full marks by giving a clear explanation what is meant by 'enrichment' as a job design technique. Many responses mentioned that enrichment involves a greater variety of tasks to perform which leads to the employee acquiring new skills. Weaker responses often just identified new skills without any further explanation. Some responses outlined either job rotation or job enlargement which was not creditworthy.

- (b) The marks achieved for this question covered the full range of the mark scheme. Stronger responses outlined the sample, some procedural details and a result of the study on workplace sabotage by Giacalone and Rosenfeld. Many responses outlined both the reasons and types of sabotage. Weaker responses often gave either the reasons or types of sabotage with no other details given. There were some incorrect responses which outlined that the study asked the participants to state the sabotage that they had done at work which was not creditworthy.
- (c) The responses to this question covered the full range of the mark scheme. Stronger responses outlined two weaknesses of the Giacalone and Rosenfeld study and gave a clear example from the study to explain their point. Common weaknesses included generalisability, socially desirable responses and weaknesses of quantitative data. Weaker responses often identified the weakness and either gave no explanation for the weakness or a very vague explanation, such as stating that the self-report could produce a socially desirable response with the participant saying what they thought the researcher wanted to hear.

Question 8

- (a) There were some good, detailed responses to this question with clear details given of leadership styles including outlines of contingency theory, situational leadership and styles of leader behaviour theories. The strongest descriptions were of Hersey and Blanchard situational leadership theory and Muczyk and Reimann's styles of leader behaviour. Contingency theory was sometimes correctly described but it had the most common confusion with responses suggesting the leader needs to adapt to the situation rather than the organisation finding the best leader (either task or relationship focused) for the situation. Weaker responses frequently outlined one of the bullet points from the section and therefore achieved in the lower mark bands. There were also some responses that gave anecdotal responses outlining the sorts of things leaders do in their jobs without any reference to theories.
- (b) The vast majority of responses did the named issue of practical applications and were able to explain how the theories of leadership styles outlined in **part (a)** could help organisations to improve performance, motivation, etc. of their employees. Few were able to do analysis to explain why these theories might be difficult to apply in organisations or in every type of organisation. Common evaluation issues also included generalisability, reductionism versus holism, and determinism versus free-will. Weaker responses often continued the description from **part (a)** and did not provide any evaluative points which was not creditworthy.

PSYCHOLOGY

Paper 9990/32 Specialist Options: Theory

Key messages

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

Candidates showed knowledge of the terminology, theories, studies and disorders identified in the syllabus as most were able to identify and/or define the terms/concepts given in these types of questions. Some of the responses showed a good level of understanding, however, some responses were either brief or anecdotal. It would be helpful to create a glossary of key terms, revision of terminology/theories using flash cards and class quizzes on terminology/theories. 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 (such as a self-report) or treatment 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 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: characteristics of anxiety disorders: measures: the blood injection phobia inventory (BIPI) to help the candidate identify which part of the syllabus the question is referring to.

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. 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 only briefly discussed the second issue. In addition, many of the responses were general and not specific to the study, theory or technique(s) named in the question. To improve, responses should give specific examples to support their point. As mentioned for the odd question **part (b)**, the candidate should make a flashcard / revision notes to help prepare for these questions, and could include in this strengths and weaknesses of the theory, study or technique. Candidates can also practice making comparisons between theories and techniques such as treatments and self-reports so that they can learn the skill of comparison during the course to prepare for these type of questions (as well as the analysis requirement of **2(b), 4(b), 6(b) and 8(b)**).

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

This question will always come from one of the bullet points in the syllabus. Candidates could describe the three (or four) studies, theories, characteristics/explanations of disorders or techniques identified in the specification under the appropriate bullet point. For this exam, the majority of the responses focussed on the correct topic area of the syllabus but sometimes gave brief and/or inaccurate descriptions. It is possible for the responses to achieve full marks by describing at least two of the studies, theories, characteristics/explanations of disorders or techniques and this would need to be a very detailed description. It is also important that the descriptions are linked to the topic area named in the syllabus. It could be useful for candidates to create revision notes with the title of each bullet point as the header in their notes.

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

This question will always ask the candidate to evaluate the theories, studies, characteristics/explanations of disorders and/or techniques described in **part (a)** of the question. The response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. However, most responses that evaluated using two issues in this exam, achieved in the lower bands due to the response being superficial and often with little analysis. Some responses that considered three issues tended to achieve higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the theories, studies, disorders and techniques described in **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. It was common in responses for candidates to identify a comparison without any explanation or examples given to explain the comparison point. A conclusion at the end of each issue would be helpful to show excellent understanding of the issue under discussion. This conclusion could summarise the analysis points that have been made for that evaluation issue or debate. To achieve the requirements of the level 3 and 4 band descriptors it would be best if the response was structured by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure the answer covers this requirement of the question.

The majority of the responses covered other issues in addition to the one named in the question. Many of the answers were structured by study/theory/technique rather than by the issue, which often led to the response being quite superficial and repetitive. It was common for responses to identify an evaluation issue and simply state whether the study, theory etc. supported the issue identified. For example, stating that a study had good ecological validity because it was done in the field without any specific example to support the discussion. Candidates should be aware that this question is worth 10 marks and include an appropriate amount of information.

General comments

The marks achieved by candidates achieved across the full range of the mark band. Many candidates were well prepared for the exam and showed good knowledge, understanding and evaluation throughout their responses. Some candidates showed limited knowledge and understanding with brief and/or superficial responses. There was often evidence of limited evaluation skills in these responses.

Time management for this paper was good for the majority candidates and most attempted all questions that were required. Some 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 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 achieved full marks. Good, concise responses often identified that a blood phobia is an irrational and/or persistent fear of blood. Many responses identified fear of needles/injections, and also the effects of being exposed to blood such as fainting. Weaker responses often just stated that it was simply to have 'fear of blood', which was not addressing the question sufficiently. Some responses were too detailed which did gain full marks but left less time to answer other questions on the paper.

- (b) Strong responses to this question gave a clear description of a number of features of the blood injection phobia inventory (BIPI). Common responses included reference to the number of items, the 4 point rating scale, specific examples of stimulus content/phobic responses, psychometric test and a measure of the severity of blood phobia. Weaker responses often just identified that the BIPI measures blood phobia and sometimes gave inaccurate features such as a 5 point rating scale and 11 items. Some incorrect responses outlined that the BIPI involves injecting a person or exposing them to blood and observing their reaction which was not creditworthy.
- (c) The responses to this question covered the full range of the mark scheme. Stronger responses identified a similarity and a difference between the BIPI and the Generalised Anxiety Disorder assessment (GAD-7) and gave an example from each self-report to explain the comparison. Most responses could identify a similarity including psychometric tests, having the same rating/quantitative data collected and measuring the severity of the disorder. Differences identified included measuring different disorders and number of items in the self-reports. The best difference identified and then explained which often achieved level 3 was that the GAD-7 is often used as a screening test which could lead to treatment or referral to a specialist whereas the BIPI can be used for a formal diagnosis. Weaker responses often just identified the comparison point with no example or explanation. Responses that were not creditworthy usually gave incorrect information about the BIPI and GAD-7. Some of the candidates got confused that the GAD-7 is for general anxiety and instead stated that the GAD-7 measures the severity of phobias.

Question 2

- (a) Responses covered the full range of the marks available. There were many good and detailed descriptions of kleptomania, pyromania and gambling disorder. Some responses did outline case studies for these disorders, but the response often focussed on the treatment of the disorder rather than the characteristics which was not addressing the question. There were many very clear and detailed descriptions of Griffiths definitions with good knowledge shown of salience, mood modification and conflict, etc. An outline was sometimes given of the Kleptomania Symptom Assessment Scale (K-SAS) although this tended to be in less detail with some inaccuracies. It was common for responses to outline features of BIPI rather than K-SAS and a significant number of responses incorrectly described Y-BOCs which is a measure of obsessive compulsive disorder so no credit was given for this part of the response. Weaker responses often lacked depth or gave details of explanations and/or treatments which was also not creditworthy for this question.
- (b) Responses for this question covered the full range of marks available with a number achieving levels 3 and 4 due to providing good analysis when discussing the named issue of self-reports in particular. Common strengths included practical application of the K-SAS for both diagnosis and monitoring treatment and the strengths of quantitative data. A common weakness discussed included social desirability. Some responses gave very good examples of why someone with kleptomania might not give honest responses to the K-SAS. Other issues that were discussed well included practical applications and individual and situational approaches to the types and definitions of impulse control disorders and nonsubstance addictive disorder. Weaker responses often gave generic discussions of self-reports in some detail but not applying this to K-SAS. A significant minority of responses went on to describe explanations and/or treatments instead of evaluating which was not creditworthy.

Psychology and Consumer Behaviour

Question 3

- (a) Most responses were able to correctly identify two of the virtual store layouts in the study by Vrechopoulos on the interior layout of grocery stores including grid, racetrack and/or free-form. Some responses gave very detailed descriptions of the virtual store layouts which was unnecessary as the question asked for identification (i.e. to name the layouts). Full credit was awarded if the response described rather than identified the layout although it was rare for candidates to not know the name of the layout if they knew the description.
- (b) The majority of responses described at least one result from the Vrechopoulos study, if not two. The stronger responses were able to make a comparison between layouts and linked to the features of time, preference or ease of use. Some of the responses had errors with a confusion over the terminology used in the study and sometimes used the term 'preferred' layout when the result for 'usefulness' was given instead. These types of responses were not creditworthy.

- (c) The marks for the responses to this question covered the full range of the mark scheme. Common strengths included generalisability due to the sample size and the participants being from two different countries, standardisation of procedure followed, practical applications, ecological validity and strengths of quantitative data collected. Common weaknesses included generalisability to other cultures and weaknesses of quantitative data.

A particularly good response given was that results may not be generalisable to other types of online stores such as clothing or video games, which then went on to describe how people shop in very different ways while looking for clothing or a book in comparison to grocery stores where they have a list to follow. Stronger responses identified the strength and weakness and gave a clear example from the study. Weaker responses often just identified the strength and/or weakness with either no example given or a very brief explanation. Some responses incorrectly identified a weakness of ecological validity of online shopping compared to real stores or did not make it clear in the response that they understood that the study was carried out online. Ecological validity was not a creditworthy point when comparing the shopping experience to real stores as the purpose of the study was to investigate the online retail experience with no intention of generalising to shopping in real stores.

Question 4

- (a) The marks for this question covered the full range of the mark scheme. Stronger responses outlined choice heuristics in consumer decision-making, including availability and representativeness. There were some clear examples from consumer decision-making and the candidates were able to apply their knowledge of these concepts to purchasing behaviour. There were some very detailed descriptions of the Wansink et al.'s research with many outlining two or three of the studies done in depth. Knutson et al. was good in general, with strong responses outlining the sample, procedure and results for specific brain regions. Weaker responses often gave incorrect details of choice heuristics by stating that availability is whether products are available in shops which was not creditworthy. Wansink et al.'s study would either be outlined in less depth or the studies would be outlined in a confused manner, often mixing up the details of the four studies done by Wansink et al. There were some level 1 responses where the candidate often gave an anecdotal outline of factors that could impact consumer decision-making.

- (b) There were some good responses to this question. These were often able to evaluate the named issue of generalisability and were able to use the studies from **part (a)** of their answer to explain how each has both positive and negative points that can be raised about the samples used, as well as the generalisability of the studies to everyday life.

Other common evaluation issues included practical applications, ecological validity and strengths and weaknesses of quantitative data. Weaker responses sometimes covered the named issue but often just identified the sample used in the study and then simply stated that it either was or was not generalisable. Some responses outlined a number of issues but in a superficial manner, simply stating that issues did or did not apply to the studies/concepts described in **part (a)** but not why. These types of responses often achieved in the level 1 mark band. In addition, there were some weak attempts to use reductionism versus holism or individual and situational explanations, but this was very rarely successful.

Psychology and Health

Question 5

- (a) There were two types of responses to this question. The first, often achieving full marks, outlined either the cost-benefit analysis or rational non-adherence and then gave an example of why this would lead to non-adherence to medical advice. Weaker responses sometimes identified the cost of treatment without explaining that this would be compared to the benefit and if the cost(s) were higher the patient would not adhere. The second type of response was anecdotal, outlining a variety of reasons patients might not adhere, such as the practitioner using too much jargon so the patient is not clear what the health advice is, lack of motivation or the patient believing their medical condition is not serious.

- (b) The marks achieved for this question covered the full range of the mark scheme. Stronger responses outlined the sample, some procedural details (including how the Funhaler works) and a result of the study on improving adherence by Watt et al. Weaker responses had fewer details of the study and frequently just described how the Funhaler worked rather than how it was used in the study. A significant number of responses stated that there were two groups of children used in the study with one using the traditional inhaler and the other the Funhaler rather than a repeated measures design where the children used the traditional inhaler for the first week and then the Funhaler one week later. In addition, a number of responses suggested that adherence was measured using a biochemical test or weighing the amount of medication left in the Funhaler rather than describing the self-report technique used with the parents/guardian in the study.
- (c) Most responses identified both a strength and weakness of the study by Watt et al. Common strengths included outlining a strength of quantitative data and practical applications of helping with adherence to other types of medical treatments and ethics. Common weaknesses included generalisability to other cultures, social desirability and a weakness of quantitative data. An example of a well explained weakness of the study was that it was not teaching children about the importance of taking medication as directed which is an important life skill but was instead suggesting medication only needs to be taken when it is fun/there is an immediate reward.

Stronger responses identified the strength and weakness and gave a clear example from the study to explain their point. Weaker responses often just identified the strength/weakness without any explanation or example. Some responses identified incorrect evaluation issues such as generalisability of the study to teenagers and adults and being unethical which were not creditworthy. The purpose of the Watt et al. study was to test the design of the Funhaler to show an improvement in adherence to taking asthma medication (and using the inhaler correctly) in children. Watt et al. did not intend the Funhaler for teens or adults.

Question 6

- (a) The responses to this question covered the full range of the mark scheme. Stronger responses showed a good understanding and details of the Weinstein study on unrealistic optimism. Most outlined the first study and a few also gave details of the follow-up study. In addition, there were some detailed descriptions of the Lau et al. study on health change in adolescence. The transtheoretical model was not always covered in responses but there were some detailed descriptions of this and sometimes this had clear examples of how this could lead to health changes in behaviour. Lower level responses often described the Weinstein study very briefly or simply stated that the participants had unrealistic optimism about their future with some examples (often not from the study). Weaker responses sometimes omitted the transtheoretical model or just listed the stages. Some weaker responses gave confused details of the Lau et al. study as to the final result regarding the relative influence of parents and peers.
- (b) There were some good responses to this question. Relating individual factors in changing health behaviours to the named issue of practical applications was often done in an effective manner. The stronger answers did this by using all three bullet points with many good examples of how health professionals could use the findings of the research and the transtheoretical model to help their patients enact health change. Some responses did then provide effective analysis usually by outlining how the study and/or model had a specific practical application and then explained the difficulties that a health professional/patient could have in carrying out this application. Other common issues chosen were self-reports (especially in relation to social desirability) and cultural differences. Weaker responses often named an issue and attempted to give examples and mainly just described the content from **part (a)**.

Psychology and Organisations

Question 7

- (a) There were many good, clear responses to this question with most achieving full marks by identifying the universalist theory of 'Great Man/Woman/Person' and that leaders are born and not made. A few responses identified charismatic leaders and then explained how these types of leaders have vision and lots of self-confidence. A small minority of responses did not appear to know this theory, and some identified contingency theory from **part (b)** which was not creditworthy.



- (b) The marks achieved for this question covered the full range of the mark scheme, although a significant number of responses outlined a theory about adaptive leadership where the leader adapts to the situation rather than contingency theory which was not creditworthy. Another common error was to outline a leadership theory from elsewhere in the syllabus which was not creditworthy. Common errors included outlining Hershey and Blanchard's situational leadership or Muczyk and Reimann's theory about styles of leader behaviour.

Stronger responses were able to describe contingency theory in some depth. These types of responses identified the least preferred co-worker scale and how a low score on the LPC meant the leader was task oriented and a high score indicated a relationship-oriented leader. Some responses were then able to describe that contingency theory suggests that leaders are matched with the situation rather than adapt to it. Weaker responses often just identified the LPC and some extended their description a little by outlining the two types of leaders (task and relationship).

- (c) The responses to this question often achieved a level 1 mark or did not explain a creditworthy strength and/or weaknesses. As many candidates did not know contingency theory and achieved no marks in **part (b)**, they found it difficult to explain a strength and a weakness of the theory and often explained a strength and/or a weakness that applied to a different leadership theory. Another common inappropriate response was to identify one or more issues that would apply to a study (e.g. good ecological validity or a small sample so poor generalisability due to a small sample size).

Candidates who had outlined contingency theory in **part (b)** were usually able to access marks with their response to this question. Stronger responses identified a strength and weakness and used contingency theory and/or the LPC scale as examples to back up their point.

Common points included providing an evaluation of a strength and a weakness of the LPC scale such as using the quantitative score from this scale to identify task and relationship-oriented leaders which can then be matched to the situation/task-structure of a project or a team in an organisation. Another common weakness was social desirability on the LPC scale. Lastly some responses explained a strength that the theory could be used to select appropriate leaders for a team/project in order for these to be managed well and therefore benefit the organisation.

Question 8

- (a) There were some good, detailed responses to this question with clear details given of temporal conditions of work environments. Stronger responses outlined the different types of shift rotation and this was sometimes done in depth with correct details given of both metropolitan and continental shift rotations. In addition, many responses gave clear details of the Gold et al. study on the nurses investigating shift-work and accidents. There were also some good details given of Knutson's review article on the effects of shift-work on health. Some responses were even able to outline the uncertainty around the health issues outlined by Knutson as to whether these health problems were caused by shift-work or other factors such as lifestyle or genetic predisposition to certain health problems, etc.

There was evidence that some candidates were not prepared to answer a question on this topic. Some weak responses did not know what temporal conditions of the work environment was referring to. Some of these responses outlined studies from other parts of the syllabus such as token economies and Three Mile Island accident which meant they were not addressing the question asked. In addition, a significant number of responses stated that shift work was a technique used by companies to persuade their employees to work shorter hours.

- (b) The majority of responses did the named issue of ecological validity and discussed how the Gold et al. study had good ecological validity. A small minority of responses did then provide some analysis. For example, the Gold et al. study was investigating accidents at work which is natural behaviour. However, the participant knew they were being investigated so this could lead to demand characteristics and/or socially desirable responses. Other common evaluation issues included a discussion of self-reports, practical applications and the individual and situational explanations. Weaker responses would frequently identify an issue and make an attempt to give an example, but this was often just re-describing the content of **part (a)**.

Evaluation points were not always logically compared and sometimes the alternative argument was unrelated to the point made in the response.

PSYCHOLOGY

Paper 9990/33 Specialist Options: Theory

Key messages

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

It is important that candidates are aware of the terminology, theories, and disorders identified in the syllabus as some were unable to identify and/or define the terms/concepts given in these types 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.

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

These questions could ask the candidate to describe a theory, study or treatment(s) used by psychologists that 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 and a result, 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, Bipolar and related disorders: treatment and management of depression: rational emotive behaviour therapy (Ellis, 1962) to help the candidate identify which part of the syllabus the question is referring to.

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. This question is worth 6 marks so the candidate should write a more extended answer for each issue raised. Some responses were very detailed for one issue but then only briefly discussed the second issue. In addition, many of the responses were general and not specific to the study, theory or technique(s) named in the question. To improve, responses should give specific examples to support their point. As mentioned for the odd question **part (b)**, the candidate should make a flashcard / revision notes and could include in this strengths and weaknesses of the theory, study or technique to help candidates prepare for these questions.

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

This question will always come from one of the bullet points in the syllabus. Candidates could describe the three (or four) studies, theories, explanations of disorders or techniques identified in the specification under the appropriate bullet point. For this exam, some of the answers used the incorrect topic area in the syllabus or the description was brief. It is possible for the responses to achieve full marks by describing at least two of the studies, theories, explanations of disorders or techniques and this would need to be a very detailed description. It is also important that the descriptions are linked to the topic area named in the syllabus. It could be useful for candidates to create revision notes with the title of each bullet point as the header in their notes.

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

This question will always ask the candidate to evaluate the theories, studies, explanations of disorders and/or techniques described in **part (a)** of the question. The response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. However, most responses that evaluated using two issues in this exam, achieved in the lower bands due to the response being superficial and often with little analysis. Some responses that considered three issues tended to achieve higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the theories, studies, disorders and techniques described in the **part (a)** of the answer. 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 to show excellent understanding of the issue under discussion. In order to achieve the requirements of the level 3 and 4 band descriptors, it would be best if the response was structured by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure that the answer covers this requirement of the question. Some of the responses covered other issues rather than the one named in the question. Quite a few of the answers were structured by study/theory/technique rather than by the issue, which often led the response to be quite superficial and repetitive. Some of the responses did analysis. Candidates should be aware this question is worth 10 marks and attempt to include an appropriate amount of information.

General comments

The marks achieved by candidates were across the full range of the mark band. Many candidates were well prepared for the exam and showed good knowledge, understanding and evaluation throughout their responses. Some candidates showed limited knowledge and understanding with brief and/or superficial responses. In these responses, there was often evidence of limited evaluation skills.

Time management for this paper was good for the majority candidates and most attempted all questions that were required. Some 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 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.

Comments on specific questions

Psychology and Abnormality

Question 1

- (a) There were many good responses to this question which identified a specific symptom of mania (e.g. increased energy) and a specific symptom of depression (e.g. withdrawal from social activities). Most responses were able to gain one mark for identifying that there was a high and a low mood phase or stating that bipolar disorder involves mania and depression. Some responses identified characteristics of other disorders such as delusions and hallucinations which were not creditworthy.
- (b) Strong responses to this question gave a clear description of REBT with reference to the ABC model (sometimes reference to ABCDE) and how this is used as a treatment for depression. Many responses did not link their description of the treatment to improving the symptoms of depression which limited the marks received. Weaker responses tended to just identify that the therapist challenges the irrational thoughts of a patient.
- (c) The responses to this question covered the full range of the mark scheme. Stronger responses identified a similarity and a difference between REBT and a biological treatment with clear examples from both to explain the comparison point. Common similarities included that both are effective for improving the symptoms of depression, both require a professional and both take time to be effective. Common differences include side effects and the approach taken to treat depression. Weaker responses often identified the similarity or difference without giving any examples to explain these. Common errors included stating that REBT deals with the cause of

depression or that SSRIs take effect immediately which were not creditworthy. In addition, a small number of responses compared REBT with CBT which was also not creditworthy.

Question 2

- (a) Responses varied considerably for this question and covered the full range of the marks available. Some responses highlighted how well prepared some of the candidates were for this exam whereas others showed limited knowledge of this topic. Some responses were detailed, accurate and coherent with a good use of psychological terminology. The strongest responses covered three of the four explanations in detail with some very good descriptions of the behavioural explanation (with the example of the Watson study given) as well as the biomedical/genetic (with the example of the Ost study). Weaker responses often gave more limited descriptions that lacked reference to terminology or clear details of the explanations and/or studies. It was common for the psychodynamic explanation to cover the whole of the Oedipus complex/description of the little Hans' study without much focus on the explanation of the development of a phobia.
- (b) Many of the responses achieved in the level 1 or level 2 mark band, with a few providing clear analysis and details of the explanations of phobias to back up their evaluative points that enabled these type of responses to achieve level 3 and above. There was a tendency for responses to focus on many issues per explanation/study rather than applying the issue to the different explanations/studies. A few responses did effectively discuss the named issue of nature versus nurture with some clear analysis. Common evaluation issues included determinism versus free-will, reductionism versus holism and application to everyday life. Weaker responses tended to identify which side of a debate the explanation was on (e.g. genetic is nature) without any explanation given for why this is the case.

Psychology and Consumer Behaviour

Question 3

- (a) Many of the responses achieved full marks by identifying two of the Cs from Lauterborn's 4 Cs marketing mix model of advertising. Responses did cover all four of the Cs including cost, convenience, communication and customer wants and/or needs. A common error was to state 'customer' on its own which was not creditworthy.
- (b) There were some clear and detailed responses describing the procedure and one result from the study by Auty and Lewis (2004) on product placement in films. There were many good descriptions of the two groups used in the study and how the procedure varied for these participants. Some responses gave details of the prompts used, and also a specific result for the study. Weaker responses often identified that the children watched the 'Home Alone' film and sometimes stated that the children showed a preference for Pepsi when exposed to this during the film but knew few specific details of either the procedure or a result.
- (c) The marks for the responses to this question covered the full range of the mark scheme. Common strengths included generalisability due to two age groups of participants, strengths of quantitative data collected, practical applications of product placement in films and how the study met ethical guidelines. Common weaknesses included generalisability and weaknesses of quantitative data. Stronger responses identified the strength and weakness and gave a clear example from the study. Weaker responses often just identified the strength and/or weakness with either no example given or a very brief explanation. Some responses incorrectly identified a weakness of ethics for the study which was not creditworthy as informed consent was given by the parent/guardian.

Question 4

- (a) The marks for this question covered the full range of the mark scheme. There was a number of good and detailed descriptions of the studies by North et al., Guéguen et al., and Woods et al. Weaker responses often gave fewer details of the studies with a general outline and conclusion given. Some responses gave muddled descriptions for the Woods et al. study.

- (b) Those responses that achieved in the higher mark bands for **part (a)**, produced strong answers to this question with a good outline of the strengths and weaknesses of the experimental method as used in the studies. Generalisability, ecological validity and applications to everyday life were other commonly chosen issues. Weaker responses often attempted a large number of evaluation issues without discussing any of them in depth or giving any examples from the **part (a)** of the response to back up their points. These type of responses often achieved level 1.

Psychology and Health

Question 5

- (a) The vast majority of responses to this question achieved one mark by identifying a characteristic of Type A personality, such as being competitive or controlling. Most responses did not explain why this characteristic could lead to a stress-related illness and often simply stated that being Type A led to experiencing more stress.
- (b) There were many strong responses to this question with many giving a detailed description of stress inoculation training (SIT) as a technique for preventing stress developed by Meichenbaum. Strong responses briefly outlined the three phases of the therapy (conceptualisation, skills acquisition and rehearsal and application and follow through) and linked this to reducing stress. Weaker responses often gave a vague description of therapy that could reduce stress that identified or indicated one of phases or identified relaxation to achieve one mark.
- (c) Most responses identified both a strength and weakness of SIT. The most common strength was effectiveness/developing the skills needed to reduce stress in the future when therapy has finished. Common weaknesses included time/cost and the level of commitment/motivation required from the client. Stronger responses identified the strength and weakness and gave a clear explanation with examples to explain their point. Weaker responses often just identified the strength/weakness without any explanation or example. Some responses identified incorrect evaluation issues such as ecological validity and self-reports which were not creditworthy.

Question 6

- (a) The responses to this question covered the full range of the mark scheme. Stronger responses showed a good understanding of biochemical, psychological and alternative techniques to manage and treat pain. These responses were able to give an outline of how the treatment worked and why it led to a reduction in pain. Many referred to the gate control theory to explain how pain is reduced. Weaker responses often identified the techniques, frequently just identifying medication without naming anything specific, and no indication was given of how these treatments would manage and/or control pain. Stronger responses were given for the psychological techniques where candidates often had more detailed knowledge of these strategies.
- (b) There were some good responses to this question. These responses were often able to evaluate the named issue of determinism versus free-will, and were able to use the treatments from **part (a)** of their answer to explain how each treatment either led to a deterministic reduction in pain and/or the patient enacted their free-will to manage their pain. Stronger responses were able to argue that all of the treatments partly involve the free-will of the patient and why. Other common evaluation issues included cost, time and side effects of medication. Weaker responses sometimes covered the named issue but often just identified whether the treatment was deterministic or down to the free-will of the patient with no example or explanation to back up the point made. Some responses outlined a number of issues but in a superficial manner, simply stating that issues did or did not apply to treatments but not why. These types of responses often achieved in the level 1 mark band.

Psychology and Organisations

Question 7

- (a) There were some good responses to this question with some achieving full marks by giving a clear explanation what is meant by 'enrichment' as a job design technique. Many responses mentioned that enrichment involves a greater variety of tasks to perform which leads to the employee acquiring new skills. Weaker responses often just identified new skills without any further explanation. Some responses outlined either job rotation or job enlargement which was not creditworthy.

- (b) The marks achieved for this question covered the full range of the mark scheme. Stronger responses outlined the sample, some procedural details and a result of the study on workplace sabotage by Giacalone and Rosenfeld. Many responses outlined both the reasons and types of sabotage. Weaker responses often gave either the reasons or types of sabotage with no other details given. There were some incorrect responses which outlined that the study asked the participants to state the sabotage that they had done at work which was not creditworthy.
- (c) The responses to this question covered the full range of the mark scheme. Stronger responses outlined two weaknesses of the Giacalone and Rosenfeld study and gave a clear example from the study to explain their point. Common weaknesses included generalisability, socially desirable responses and weaknesses of quantitative data. Weaker responses often identified the weakness and either gave no explanation for the weakness or a very vague explanation, such as stating that the self-report could produce a socially desirable response with the participant saying what they thought the researcher wanted to hear.

Question 8

- (a) There were some good, detailed responses to this question with clear details given of leadership styles including outlines of contingency theory, situational leadership and styles of leader behaviour theories. The strongest descriptions were of Hersey and Blanchard situational leadership theory and Muczyk and Reimann's styles of leader behaviour. Contingency theory was sometimes correctly described but it had the most common confusion with responses suggesting the leader needs to adapt to the situation rather than the organisation finding the best leader (either task or relationship focused) for the situation. Weaker responses frequently outlined one of the bullet points from the section and therefore achieved in the lower mark bands. There were also some responses that gave anecdotal responses outlining the sorts of things leaders do in their jobs without any reference to theories.
- (b) The vast majority of responses did the named issue of practical applications and were able to explain how the theories of leadership styles outlined in **part (a)** could help organisations to improve performance, motivation, etc. of their employees. Few were able to do analysis to explain why these theories might be difficult to apply in organisations or in every type of organisation. Common evaluation issues also included generalisability, reductionism versus holism, and determinism versus free-will. Weaker responses often continued the description from **part (a)** and did not provide any evaluative points which was not creditworthy.

PSYCHOLOGY

Paper 9990/41 Specialist Options: Application
--

Key messages

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

General comments

A few candidates answering questions from one option only. Some candidates answered questions from three and even four options. Of those correctly selecting from two options, whilst answers to one option were often very good, some answers to the second option were much weaker, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

- (i) Candidates are advised to read the 'stem' of the question, the introduction or the opening words in **Section A** questions as the information provided is crucial to answering each question part that follows.
- (ii) Answers must refer to the study the question is about. Many answers provided general comments which were unrelated to the study itself.
- (iii) For question **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so restricted marks. For example, to score 1 mark, answers must include a strength (or weakness) and an example.

Candidates should not use terms without explanation, frequently answers were restricted to '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. Many candidates appeared to 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 merely repeated what had already been written. 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 read the question and its requirements carefully, 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 show the Examiner their psychological knowledge.

Section B

In many cases, candidates attempted to conduct an experiment even if the question did not require it. 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. However, it is not necessary 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. It does not require candidates to describe everything they know about that topic area, and answers that don't address the question will only achieve minimal marks. To score marks at the top end of the mark range, answers must focus on arguments both for and against the statement and use appropriate evidence to support the arguments. At the very top of the mark range, answers should show awareness of wider issues and evidence that is relevant.

Comments on specific questions

Section A

Question 1

- (a) Many answers were awarded full marks for explaining the terms aversive and imagery. Some candidates used the example from the study by Glover where the participant imagined nausea and vomiting which confirmed the awarding of full marks. Some candidates wrote that imagery involves showing pictures to participants which is incorrect because imagery is using imagination rather than actual pictures.
- (b) Treatment for kleptomania can be given in a number of ways and one way is face-to-face with a therapist. Many candidates provided two good strengths which were creditworthy. However, the strengths must be related to the specific topic of the question (in this instance kleptomania) to allow full credit to be given. If the strength is not related (to kleptomania) the strength could relate to any topic area that uses a therapist for treatment. Candidates must apply their general psychological knowledge to specific syllabus topics.
- (c) Many candidates were awarded 2 marks for making appropriate suggestions to explain why biochemical treatments might not be effective, such as medications may not be taken; that they may be addictive; that they might reduce symptoms rather than the cause. However, very few candidates were awarded the additional 2 marks because they needed to address the 'when treating kleptomania' part of the question. A simple 'there is no drug that stops a person from stealing' would be sufficient.

- (d) Most candidates had few problems with stating why generalisations can and cannot be made from a case study. A few candidates knew the case study by Glover and applied it well; others did not relate their discussion to the Glover study at all. **Section A part (d)** answers for all options require a conclusion. A conclusion is not a summary, and so candidates summarising what had already been written scored no marks. A conclusion is a 'decision reached by reasoning' and only answers addressing this were awarded the 1 available mark for a conclusion.

Question 2

- (a) Candidates had to suggest a sampling technique that was used in the study. The correct answer was opportunity sampling. Use of the words in the stem of the question such as 'students ... recruited from around the campus' could only result in an opportunity sample being applied. Random sampling and volunteer sampling were also suggested by some candidates.
- (b) (i) There were several possible strengths of the sample, including that it was large, had a wide age range and that there was a nearly equal male/female balance. Suggesting any one of these scored 1 mark, and those candidates who supported their suggestion with an example from the study (240 participants, 19-35 years, 52 per cent to 48 per cent) were awarded the additional 1 mark.
- (ii) There were several possible weaknesses of the sample. It could be said that the sample was small (only 240 participants), that the age range was restricted (no-one above 35 years), that the participants were all students, or that the students may not be very familiar with ordering food from a menu.
- (c) (i) This question invited candidates to identify the two (of the four) versions of the menu used in the Dayan and Bar-Hillel (2011) study. Most answers were incorrect. Those knowing the study were awarded 2 marks for identifying two of 'baseline menu', 'mirror menu', 'inside-out base' or 'inside-out mirror'.
- (ii) Most candidates could explain why participants are randomly allocated to conditions of an experiment and scored 1 mark. However, most could not go further and apply their knowledge to this study. A simple reference to the conditions of the IV, 'baseline menu', 'mirror menu', 'inside-out base' or 'inside-out mirror' would have been sufficient.
- (d) Many answers included two strengths and two weaknesses of using a laboratory experiment, but often only scored partial marks because answers were not related to the study of menu item order as the question required. A comment such as 'studying menu item order in a laboratory would be artificial because participants would not be ordering food items which they would then eat, unlike in a real restaurant' would have been sufficient. Many answers had no appropriate conclusion (see **Question 1(d)** above).

Question 3

- (a) Most candidates could correctly explain the term structured interview without confusion with other types. Answers such as 'asking the same questions to every participant (1 mark) in the same order (1 mark)' was sufficient for full marks. A small number of candidates incorrectly believed that quantitative data was a feature of a structured interview and scored 0 marks.
- (b) Answers to this question part were most frequently awarded partial marks because responses did not address the 'as used in this study' part of the question. In this instance this meant that candidates were being awarded 2 marks out of 4 rather than 4 out of 4. Most candidates could give a strength and a weakness of asking closed questions but could not relate this to the McKinstry and Wang study.
- (c) (i) and (ii) These two question parts were related to each other. **Part (i)** asked how data could be analysed, the most logical answer being to calculate the mean for each style of dress. **Part (ii)** wanted a suggestion about how this data could be shown on a graph, and if the mean scores had been calculated in (i) these could be plotted on a bar chart in **part (ii)**. However, answers to **part (i)** were weak. For (ii) many candidates could correctly suggest using a bar chart, often stating 'Bar chart plotting mean scores for each style of clothing' (1 mark) and 'X axis bar for each style of clothing; Y axis for level of confidence' (2 marks) and this meant that they knew the mean score had to be calculated.

- (d) Generally, the strengths and weaknesses suggested in response to this question were weak. Candidates should be mindful that this is an applied paper where thinking and applying is required in addition to pre-prepared knowledge. To achieve marks, answers can be quite simple. For example, stating that 'photographs are standardised, remaining the same for all participants' (1 mark) 'where the expression and dress of the doctor remain constant' (2 marks).

Question 4

- (a) A few responses were very strong with their logic in working out the most likely method used in the Cowpe (1989) study which was often concluded to be a field experiment. The logic applied was that the advertisements in different regions was the IV. A small number of candidates suggested that it was a quasi-experiment, where there is control over the procedure but not participants, and this answer was also given credit. Answers suggesting a laboratory experiment were incorrect, as were those suggesting that it was a natural experiment.
- (b) A number of candidates were awarded 2 marks for correctly identifying the two strategies and some candidates were awarded 4 marks for both identifying the strategies and explaining how they were used in the study. The two were: a prevention strategy (1 mark) telling people how to avoid a chip pan fire from starting (+1 mark) and a containment strategy (+1 mark) educating people on correct and incorrect procedure to follow if a chip pan fire starts (+1 mark). Credit was also awarded to those candidates suggesting that the two strategies were providing information and fear-arousal.
- (c) (i) Some candidates scored 2 marks for stating 'the number of chip pan fires was measured through fire brigade statistics which showed a 7 per cent – 12 per cent reduction over the 12-month duration of the study'. Others scored 1 mark for 'fire brigade statistics'.
- (ii) The strength of this data is that it is objective (1 mark) because either the fire brigade attends a fire, or they do not (2 marks). There is no subjectivity or bias evident at all.
- (d) Whilst the 'using television' component was always addressed, the 'to promote safety in organisations' component of the question was not, and so candidates were restricted to partial marks.

Section B

Question 5

- (a) Nearly all candidates chose to design a laboratory experiment, even though the question allowed an open choice. Answers at the top end of the mark range included an extensive range of both specific and general features. Each feature was explained, showing that the candidate clearly understood how the feature applied to testing the effectiveness of biochemical treatments. Top answers chose to focus on gambling (as their impulse control disorder) and nalmefene was used as their biochemical treatment. Answers at the bottom end were not clear about what biochemical treatment they were using or which ICD they were designing a study for.

- (b) The psychological evidence of the study by Grant et al. (2008), who used the biochemical treatment of nalmefene, was appropriate for those candidates choosing gambling as their ICD, but less appropriate for candidates choosing to design a study involving pyromania or kleptomania. Candidates are encouraged to think through their design before beginning to write their answer. Overall methodological evidence was very good, although to maximise marks candidates should explain two or three features of their chosen design rather than provide a long list of unexplained features that are not related to their design.

Question 6

- (a) Not all candidates were able to provide a coherent design to answer this question. Many started with a field experiment, moved into an observation and ended with a questionnaire. This meant that no method was done in sufficient detail. This also meant that procedures were often confusing. One way to answer this question would be to have participants find a food item, tracking their route using observation; another way would be to use a questionnaire to ask about the accuracy and type of memory for locating food items.
- (b) Very few candidates wrote about the psychological knowledge of Mackay and Olshavsky (1975) who studied cognitive maps in retail locations. Instead work from other studies was made to fit. Whilst this was occasionally done well, often it was not. Methodologically it is always advised to explain the reasons underlying two or three design features, but it is not advisable to do these from a number of different methods.

Question 7

- (a) Candidates provided excellent answers based on doctor-centred and patient-centred strategies. However, what was largely absent from most answers was the application of this when making a correct or incorrect diagnosis. Many candidates simply looked at the preference of a patient after being presented with a strategy. Some plans became confused when using a repeated measures design where a patient would be exposed to both strategies (the IV) but then the DV measured whether the doctor provided a correct or incorrect diagnosis meaning that the strategy used with the patient became irrelevant.
- (b) The studies by Byrne and Long (1976) and Savage and Armstrong (1990) featured as psychological knowledge and many candidates described one or both studies accurately. However, marks were often restricted because answers did not explain how either of these studies had informed their design. Although the syllabus has the sub-section 'practitioner diagnosis: type I and type II errors', this was absent from most answers. Methodologically answers were good, but some candidates would benefit from ensuring that explanations for their design decisions are provided rather than just simple statements.

Question 8

- (a) Most candidates used a questionnaire to ask participants about the effectiveness of different types of non-monetary reward. There were some very strong answers, but also weaker responses. There were two main differences between strong and weaker answers. Some candidates included a full range of features appropriate to questionnaires, others stated just 'a questionnaire'. Some included appropriate questions showing good knowledge of different non-monetary rewards, whilst others simply asked: 'which one do you prefer'.
- (b) Following on from (a) above, psychological knowledge was in many instances very strong, and referred to praise, respect, recognition, empowerment and sometimes a sense of belonging. Methodologically, some candidates provided good explanations for their choice of question type, for example, opting for closed questions to gather quantitative data which they could analyse statistically. Some candidates appeared to suggest that if they gather quantitative data, they automatically need to also gather qualitative data, which is not the case.

Section C

Question 9

Answers scoring the highest marks were those which were organized and answered the question set. This meant that two halves of an answer were evident. Reasons supporting the view that psychometric measures are useful were followed by reasons why such measures may not be useful. Measures such as the Maudsley Obsessive-Compulsive Inventory (MOCI) and Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) featured in strong answers. Some candidates wrote descriptive essays which did not address the question set and could not be credited.

Question 10

Many candidates appeared to misunderstand the nature of the study by Knutson et al. (2007) on pre-cognitive decisions, on which this question was based. This study showed that the brain has made a decision before the person becomes cognitively aware of that decision. The common assumption is that the brain thinks one thing, yet cognitively a person might make a different decision. Appropriate answers often focused on the strengths of using fMRI to understand more about how the brain works (when making consumer decisions), whilst also arguing that when shopping it is not possible to take the fMRI along when the person would make a decision.

Question 11

Answers achieving the highest marks focused their attention on the arguments for and against the use of biochemical tests to measure adherence and threw into their discussions evidence from relevant research such as that by Roth and Caron (1978). Some candidates gave descriptive essays about adherence to medical advice followed with evaluation of their description, which did not address the question set. There are no AO1 marks for this question; all the marks are AO3 and so the question must be answered appropriately.

Question 12

The most logical approach to achieve a good answer was to provide a number of similarities between the theories of Maslow and Alderfer and then to consider all the differences (or vice versa). A few candidates took this approach and scored high marks. However, many candidates merely described Maslow's hierarchy and then described Alderfer's ERG theory without attempt to extract what was similar or different, and therefore did not score high marks.

PSYCHOLOGY

Paper 9990/42 Specialist Options: Application
--

Key messages

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

General comments

A few candidates answering questions from one option only. Some candidates answered questions from three and even four options. Of those correctly selecting from two options, whilst answers to one option were often very good, some answers to the second option were weaker, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

- (i) Candidates are advised to read the 'stem' of the question, the introduction or the opening words in **Section A** questions as the information provided is crucial to answering each question part that follows.
- (ii) Answers must refer to the study the question is about. Many answers provided general comments which were unrelated to the study itself.
- (iii) For question **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so restricted marks. For example, to score 1 mark, answers must include a strength (or weakness) and an example.

Candidates should not use terms without explanation, frequently answers were restricted to '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. Many candidates appeared to 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 merely repeated what had already been written. 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 are encouraged to read the question and its requirements carefully, 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 show the Examiner their psychological knowledge.

Section B

In many cases, candidates attempted to conduct an experiment even if the question did not require it. 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. However, 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. It does not require candidates to describe everything 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 and use appropriate evidence to support the argument. At the very top of the mark range, answers should show awareness of wider issues and evidence that is relevant.

Comments on specific questions

Section A

Question 1

- (a) Many answers considered logical differences, such as the use of medication compared to no medication, or passive treatment (swallowing a pill) compared to active treatment (such as CBT). A number of candidates incorrectly focused on explanations, which they claimed are either nature of nurture, but a treatment cannot be either of these and such answers could not be credited.
- (b) Many strong responses gave two appropriate differences between medical and psychological techniques. For example, 2 marks were awarded for stating 'antipsychotics were first generation, 1950's, whereas atypical antipsychotics were second generation, 1990's'. A second difference was in relation to side effects. Some candidates only scored a partial 1 mark for stating 'antipsychotic have more side effects and atypical less' whereas candidates giving an example of what these side effects were (such as extra pyramidal symptoms and tardive dyskinesia) scored 2 marks.
- (c) All candidates scored at least 1 mark with many scoring the full 4 marks. Marks were awarded for including muscle relaxant given, shock applied laterally or bilaterally, patient convulses, patient is unconscious, then recovers. A few candidates mentioned side effects but this could not be credited because the question was how ECT is used.

- (d) Not all candidates were able to provide a range of strengths and weaknesses in relation to the use of psychological techniques to manage schizophrenia. However, three problems were apparent. Firstly, many candidates evaluated the specific studies by Paul and Lenz (1977) and Sensky (2000), referring to sample size, for example. This type of answer could not be credited. Second, some candidates presented ECT as a psychological treatment when it is not. Thirdly, some candidates identified issues, such as reductionism, but did not explain how the issue was a strength or weakness, or gave vague statements, such as 'psychological treatments are useful for everyday life'. Similarly, 'psychological treatments are free-will because patients can choose' without further elaboration. Patients can apply 'free-will' in deciding whether to take prescribed medication or not.

Question 2

- (a) According to Porublev et al. (2009) gifts can be wrapped traditionally or non-traditionally and candidates including either of these options scored 1 mark, with an explanation of what this would involve scoring the second mark. Some candidates described a wrapped gift (as they would wrap it), and although this scored 1 mark, some knowledge of the Porublev et al. study was needed for full marks.
- (b) Many answers were awarded full marks. Some candidates knew the study and knew why a semi-structured interview was used; others did not appear to know the study or what a semi-structured interview involved. For example, there was the incorrect assumption that a semi-structured interview is automatically qualitative data. Some responses needed to address the 'in this study' part of the question.
- (c) This question invited candidates to suggest two features of the observation used in the study by Porublev et al. Some candidates were unable to suggest any features; others suggested two features from the 'standard four' (covert/overt; participant/non-participant; structured/unstructured; controlled/naturalistic) but did not relate either to the study (the question stated 'used in this study'). Some candidates provided excellent answers and scored 4 marks for relating two features to the Porublev et al. study.
- (d) Many answers included two strengths and two weaknesses of observations, but these were often in general terms and only occasionally related to gift wrapping preference. Many candidates assumed that an observation is automatically unethical, which is not correct. For example, we do not need the consent of every person we look at as we walk out in public, i.e. naturalistic observation of natural behaviour. If a situation is controlled (manipulated) then consent to take part in a study (be observed) would be needed.

Question 3

- (a) The term biofeedback appeared to be misunderstood by many candidates. Biofeedback is a technique where people control their own biological functions. For example, GSR measures skin conductance which changes with levels of physiological arousal. A high pitch tone indicates high levels of arousal (stress) and by relaxing, the tone can be lowered, reducing stress. It does not involve expensive equipment, like an fMRI, as many responses suggested.
- (b) Most candidates were awarded full marks for providing two findings, each supported with data appearing in fig 3.1. Other candidates were only awarded partial marks when stating, for example, 'one finding is that the variable pitch was best' with no supporting numbers or further comment.
- (c) Most candidates were awarded full marks for suggesting two techniques that could be used for managing stress. Stress inoculation training was often mentioned, as was the use of imagery and relaxation. A few candidates incorrectly reverted to the abnormality option and suggested techniques for treating impulse control disorders (such as imaginal desensitisation).

- (d) Generally the strengths and weaknesses suggested in response to this question were very poor. Some responses claimed that 'biofeedback is reductionist' with no explanation of what this meant or whether this was a strength or weakness. As biofeedback includes both physiological and cognitive components, it is not reductionist. It was often claimed that the pitch of the feedback was unethical and harmful to participants, when the volume of the sound could be turned down. Some responses also claimed that biofeedback equipment is too expensive for most people. In fact, a GSR machine is a very cheap, one-off purchase, and is far less costly than any long-term medication.

Question 4

- (a) Answers such as 'job rotation is where a job is rotated' could not be credited, likewise answers which focused on shift work. Shift work is incorrect because here the worker does the same job but at different times. Job rotation keeps the worker at the same level of responsibility; there is simply a change in task to be done which, although reduces boredom, is unlikely to improve job satisfaction.
- (b) The terms job enrichment and job enlargement were often misunderstood. Both involve more jobs/tasks for workers, and both would involve more variety. However, the differences between the two are that enlargement keeps a worker at the same level of responsibility and at the same skill level, whereas enrichment involves more responsibility and increased new skills.
- (c) (i) Many candidates scored 2 marks for correctly describing one of three correct psychometric tests used to measure worker satisfaction, which were the Minnesota Satisfaction Questionnaire (MSQ); the Quality of Work Life (QWL) evaluation scale; and the Job Description Index (JDI). A few candidates suggested giving workers a questionnaire, but a 'general' questionnaire is not a psychometric test.
- (ii) Many candidates achieved no marks because they suggested giving the workers a questionnaire or interview despite the question stating: 'other than using a self-report'. The syllabus defines self-report as a questionnaire or interview. Strong answers suggested using an observation of worker behaviour, and the gathering of attendance/absenteeism data to measure worker satisfaction was also creditworthy.
- (d) Appropriate advantages and disadvantages for interviews were included, but frequently these were not related to job satisfaction. Many candidates took a negative view of the interview process, perhaps reflecting management attitudes toward workers in their country. Measuring job satisfaction is a component of many performance appraisals which more often than not, gives a worker positive feedback and is designed to motivate them and look at ways in which they can be promoted.

Section B

Question 5

- (a) Designs were of three types: the first type included an extensive range of both specific features of an experiment (the most common method chosen) and an appropriate range of general features. Each feature was explained, showing the candidate clearly understood how the features applied to testing the effectiveness of chemical/drug treatments. The second type was similar to the first but lacked explanation. Often responses were restricted to a list of features, for example, 'my study will be double blind, independent and longitudinal'. The third type of answer lacked features, explanation, and there was often confusion regarding the design. For example, candidates were unclear whether participants were being given MAOIs, SSRIs or both; some appeared to assume that participants only needed to be given drugs for one week to measure their effectiveness.
- (b) The psychological evidence often included descriptions of how MAOIs (or SSRIs) work but with no reference about how this related to the design of the study (as is needed to score all the available marks). For methodological knowledge, many responses would improve with more thought, for example, it was suggested that the reason for using an independent design was to reduce demand characteristics, not the fact that using a repeated design would confound the whole study. To maximise marks, candidates should explain two or three features of their chosen design, rather than provide a long list of unexplained features that are not related to their design.

Question 6

- (a) The design of an experiment was required, and most candidates included an IV, DV and a range of appropriate methodological terminology. One weakness in the design of answers was that the task participants were required to complete was too complex, particularly with many of the very young age ranges used in designs. Some candidates used only children and often an age range was as low as three or four years of age – meaning that the required task would be impossible for children so young. Some participants got around this by having adults complete a ‘word’ task and children complete a ‘picture’ task, but these are different stimuli and the results would not be comparable.
- (b) Many candidates wrote about the work of Fischer et al. (1991) on brand recognition in children, which might be why many opted to study age differences in children specifically in part (a). Descriptions of the study were good, and in some instances the description was linked to the candidate’s design as it should be. Methodologically, a good range of points was often made, though two or three points with explanation is better than a long, bullet-point list.

Question 7

- (a) There were some strong answers to this question, gender differences in hypochondriasis, but most candidates focused on whether a person thought they had hypochondriasis or not. For example, many designs used a questionnaire to ask participants in a waiting room whether they had hypochondriasis or not. This would not be helpful given that less than 1% of the population have it. A better design could have been to give a questionnaire to doctors asking about people with hypochondriasis they have diagnosed or treated, and their gender.
- (b) The work of Barlow and Durand (1995) should have featured here, but many responses had little or no focus on hypochondriasis. The implication of this was that **part (a)** questionnaires were poorly designed (as mentioned above). Some candidates confused hypochondriasis with Munchausen Syndrome.

Question 8

- (a) A wide range of different methods were used to answer this question, including observations, questionnaires, interviews and experiments, and some candidates opted for several of these rather than focusing on just one. Very few candidates opted for a natural experiment with IV as days of week and DV as the number of absences. Some candidates interviewed participants to ask about the reason for absence, however the question did not require this to be included.
- (b) Very few candidates picked up that recording number of workers absent from work on specific days is objective data, a major strength when compared to subjective self-reports. Methodologically, days of the week cannot be manipulated and neither could workers be in conditions of an IV to be either absent or not. In terms of sample, many responses compared different departments within an organisation (which became an IV) rather than including every worker within an organisation. Psychological knowledge often involved descriptions of absenteeism by Blau and Boal (1987), but often there was nothing more than a description of the different types of absence rather than application of this knowledge to the design of the study.

Section C

Question 9

Answers achieving the highest marks outlined why the nature-nurture debate is relevant and then considered the arguments why the debate is not relevant. Explanations such as genetic and neurochemical, and cognitive and learned helplessness were common features. Most responses focused on explanations, as the question required, but some responses suggested that the debate was not relevant and began considering treatments. A few candidates did not focus on the question and addressed determinism and free-will or reductionism and holism. A few responses evaluated individual explanations of depression and often made no reference to nature or nurture.

Question 10

Many candidates correctly focused on the appropriate syllabus section of 'advertising media (e.g. television)' and wrote some very strong answers. Strongest answers compared the effectiveness of television advertising to radio, social media, or newspapers/posters. However, many responses incorrectly interpreted the question as a general advertising question and included the various 'advertising applications', such as brand recognition, or focused on product placement or marketing mix models.

Question 11

Many different approaches were taken when answering this question. The most logical approach was to consider the reliability and validity of psychological measures followed by a consideration of the reliability and validity of physiological measures. Whilst many candidates did this, some considered just psychological or physiological or focused on managing stress (rather than measuring) and some wrote extensively about the measures, although the question was about stress.

Question 12

This question required the candidates to focus on whether Belbin's team roles are the perfect way to categorise people into team roles. Some answers were restricted to a description of the three main types and the nine team roles rather than a discussion. Some candidates considered alternative theories for groups and their roles, such as that proposed by Tuckman (1965), which was an appropriate alternative to Belbin and fully creditworthy. There were also responses that talked about leadership theories, which did not address the question and were not creditworthy.

PSYCHOLOGY

Paper 9990/43 Specialist Options: Application
--

Key messages

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

General comments

A few candidates answering questions from one option only. Some candidates answered questions from three and even four options. Of those correctly selecting from two options, whilst answers to one option were often very good, some answers to the second option were much weaker, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

- (i) Candidates are advised to read the 'stem' of the question, the introduction or the opening words in **Section A** questions as the information provided is crucial to answering each question part that follows.
- (ii) Answers must refer to the study the question is about. Many answers provided general comments which were unrelated to the study itself.
- (iii) For question **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so restricted marks. For example, to score 1 mark, answers must include a strength (or weakness) and an example.

Candidates should not use terms without explanation, frequently answers were restricted to '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. Many candidates appeared to 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 merely repeated what had already been written. 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 read the question and its requirements carefully, 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 show the Examiner their psychological knowledge.

Section B

In many cases, candidates attempted to conduct an experiment even if the question did not require it. 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. However, it is not necessary 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. It does not require candidates to describe everything they know about that topic area, and answers that don't address the question will only achieve minimal marks. To score marks at the top end of the mark range, answers must focus on arguments both for and against the statement and use appropriate evidence to support the arguments. At the very top of the mark range, answers should show awareness of wider issues and evidence that is relevant.

Comments on specific questions

Section A

Question 1

- (a) Many answers were awarded full marks for explaining the terms aversive and imagery. Some candidates used the example from the study by Glover where the participant imagined nausea and vomiting which confirmed the awarding of full marks. Some candidates wrote that imagery involves showing pictures to participants which is incorrect because imagery is using imagination rather than actual pictures.
- (b) Treatment for kleptomania can be given in a number of ways and one way is face-to-face with a therapist. Many candidates provided two good strengths which were creditworthy. However, the strengths must be related to the specific topic of the question (in this instance kleptomania) to allow full credit to be given. If the strength is not related (to kleptomania) the strength could relate to any topic area that uses a therapist for treatment. Candidates must apply their general psychological knowledge to specific syllabus topics.
- (c) Many candidates were awarded 2 marks for making appropriate suggestions to explain why biochemical treatments might not be effective, such as medications may not be taken; that they may be addictive; that they might reduce symptoms rather than the cause. However, very few candidates were awarded the additional 2 marks because they needed to address the 'when treating kleptomania' part of the question. A simple 'there is no drug that stops a person from stealing' would be sufficient.

- (d) Most candidates had few problems with stating why generalisations can and cannot be made from a case study. A few candidates knew the case study by Glover and applied it well; others did not relate their discussion to the Glover study at all. **Section A part (d)** answers for all options require a conclusion. A conclusion is not a summary, and so candidates summarising what had already been written scored no marks. A conclusion is a 'decision reached by reasoning' and only answers addressing this were awarded the 1 available mark for a conclusion.

Question 2

- (a) Candidates had to suggest a sampling technique that was used in the study. The correct answer was opportunity sampling. Use of the words in the stem of the question such as 'students ... recruited from around the campus' could only result in an opportunity sample being applied. Random sampling and volunteer sampling were also suggested by some candidates.
- (b) (i) There were several possible strengths of the sample, including that it was large, had a wide age range and that there was a nearly equal male/female balance. Suggesting any one of these scored 1 mark, and those candidates who supported their suggestion with an example from the study (240 participants, 19-35 years, 52 per cent to 48 per cent) were awarded the additional 1 mark.
- (ii) There were several possible weaknesses of the sample. It could be said that the sample was small (only 240 participants), that the age range was restricted (no-one above 35 years), that the participants were all students, or that the students may not be very familiar with ordering food from a menu.
- (c) (i) This question invited candidates to identify the two (of the four) versions of the menu used in the Dayan and Bar-Hillel (2011) study. Most answers were incorrect. Those knowing the study were awarded 2 marks for identifying two of 'baseline menu', 'mirror menu', 'inside-out base' or 'inside-out mirror'.
- (ii) Most candidates could explain why participants are randomly allocated to conditions of an experiment and scored 1 mark. However, most could not go further and apply their knowledge to this study. A simple reference to the conditions of the IV, 'baseline menu', 'mirror menu', 'inside-out base' or 'inside-out mirror' would have been sufficient.
- (d) Many answers included two strengths and two weaknesses of using a laboratory experiment, but often only scored partial marks because answers were not related to the study of menu item order as the question required. A comment such as 'studying menu item order in a laboratory would be artificial because participants would not be ordering food items which they would then eat, unlike in a real restaurant' would have been sufficient. Many answers had no appropriate conclusion (see **Question 1(d)** above).

Question 3

- (a) Most candidates could correctly explain the term structured interview without confusion with other types. Answers such as 'asking the same questions to every participant (1 mark) in the same order (1 mark)' was sufficient for full marks. A small number of candidates incorrectly believed that quantitative data was a feature of a structured interview and scored 0 marks.
- (b) Answers to this question part were most frequently awarded partial marks because responses did not address the 'as used in this study' part of the question. In this instance this meant that candidates were being awarded 2 marks out of 4 rather than 4 out of 4. Most candidates could give a strength and a weakness of asking closed questions but could not relate this to the McKinstry and Wang study.
- (c) (i) and (ii) These two question parts were related to each other. **Part (i)** asked how data could be analysed, the most logical answer being to calculate the mean for each style of dress. **Part (ii)** wanted a suggestion about how this data could be shown on a graph, and if the mean scores had been calculated in (i) these could be plotted on a bar chart in **part (ii)**. However, answers to **part (i)** were weak. For (ii) many candidates could correctly suggest using a bar chart, often stating 'Bar chart plotting mean scores for each style of clothing' (1 mark) and 'X axis bar for each style of clothing; Y axis for level of confidence' (2 marks) and this meant that they knew the mean score had to be calculated.

- (d) Generally, the strengths and weaknesses suggested in response to this question were weak. Candidates should be mindful that this is an applied paper where thinking and applying is required in addition to pre-prepared knowledge. To achieve marks, answers can be quite simple. For example, stating that 'photographs are standardised, remaining the same for all participants' (1 mark) 'where the expression and dress of the doctor remain constant' (2 marks).

Question 4

- (a) A few responses were very strong with their logic in working out the most likely method used in the Cowpe (1989) study which was often concluded to be a field experiment. The logic applied was that the advertisements in different regions was the IV. A small number of candidates suggested that it was a quasi-experiment, where there is control over the procedure but not participants, and this answer was also given credit. Answers suggesting a laboratory experiment were incorrect, as were those suggesting that it was a natural experiment.
- (b) A number of candidates were awarded 2 marks for correctly identifying the two strategies and some candidates were awarded 4 marks for both identifying the strategies and explaining how they were used in the study. The two were: a prevention strategy (1 mark) telling people how to avoid a chip pan fire from starting (+1 mark) and a containment strategy (+1 mark) educating people on correct and incorrect procedure to follow if a chip pan fire starts (+1 mark). Credit was also awarded to those candidates suggesting that the two strategies were providing information and fear-arousal.
- (c) (i) Some candidates scored 2 marks for stating 'the number of chip pan fires was measured through fire brigade statistics which showed a 7 per cent – 12 per cent reduction over the 12-month duration of the study'. Others scored 1 mark for 'fire brigade statistics'.
- (ii) The strength of this data is that it is objective (1 mark) because either the fire brigade attends a fire, or they do not (2 marks). There is no subjectivity or bias evident at all.
- (d) Whilst the 'using television' component was always addressed, the 'to promote safety in organisations' component of the question was not, and so candidates were restricted to partial marks.

Section B

Question 5

- (a) Nearly all candidates chose to design a laboratory experiment, even though the question allowed an open choice. Answers at the top end of the mark range included an extensive range of both specific and general features. Each feature was explained, showing that the candidate clearly understood how the feature applied to testing the effectiveness of biochemical treatments. Top answers chose to focus on gambling (as their impulse control disorder) and nalmefene was used as their biochemical treatment. Answers at the bottom end were not clear about what biochemical treatment they were using or which ICD they were designing a study for.

- (b) The psychological evidence of the study by Grant et al. (2008), who used the biochemical treatment of nalmefene, was appropriate for those candidates choosing gambling as their ICD, but less appropriate for candidates choosing to design a study involving pyromania or kleptomania. Candidates are encouraged to think through their design before beginning to write their answer. Overall methodological evidence was very good, although to maximise marks candidates should explain two or three features of their chosen design rather than provide a long list of unexplained features that are not related to their design.

Question 6

- (a) Not all candidates were able to provide a coherent design to answer this question. Many started with a field experiment, moved into an observation and ended with a questionnaire. This meant that no method was done in sufficient detail. This also meant that procedures were often confusing. One way to answer this question would be to have participants find a food item, tracking their route using observation; another way would be to use a questionnaire to ask about the accuracy and type of memory for locating food items.
- (b) Very few candidates wrote about the psychological knowledge of Mackay and Olshavsky (1975) who studied cognitive maps in retail locations. Instead work from other studies was made to fit. Whilst this was occasionally done well, often it was not. Methodologically it is always advised to explain the reasons underlying two or three design features, but it is not advisable to do these from a number of different methods.

Question 7

- (a) Candidates provided excellent answers based on doctor-centred and patient-centred strategies. However, what was largely absent from most answers was the application of this when making a correct or incorrect diagnosis. Many candidates simply looked at the preference of a patient after being presented with a strategy. Some plans became confused when using a repeated measures design where a patient would be exposed to both strategies (the IV) but then the DV measured whether the doctor provided a correct or incorrect diagnosis meaning that the strategy used with the patient became irrelevant.
- (b) The studies by Byrne and Long (1976) and Savage and Armstrong (1990) featured as psychological knowledge and many candidates described one or both studies accurately. However, marks were often restricted because answers did not explain how either of these studies had informed their design. Although the syllabus has the sub-section 'practitioner diagnosis: type I and type II errors', this was absent from most answers. Methodologically answers were good, but some candidates would benefit from ensuring that explanations for their design decisions are provided rather than just simple statements.

Question 8

- (a) Most candidates used a questionnaire to ask participants about the effectiveness of different types of non-monetary reward. There were some very strong answers, but also weaker responses. There were two main differences between strong and weaker answers. Some candidates included a full range of features appropriate to questionnaires, others stated just 'a questionnaire'. Some included appropriate questions showing good knowledge of different non-monetary rewards, whilst others simply asked: 'which one do you prefer'.
- (b) Following on from (a) above, psychological knowledge was in many instances very strong, and referred to praise, respect, recognition, empowerment and sometimes a sense of belonging. Methodologically, some candidates provided good explanations for their choice of question type, for example, opting for closed questions to gather quantitative data which they could analyse statistically. Some candidates appeared to suggest that if they gather quantitative data, they automatically need to also gather qualitative data, which is not the case.

Section C

Question 9

Answers scoring the highest marks were those which were organized and answered the question set. This meant that two halves of an answer were evident. Reasons supporting the view that psychometric measures are useful were followed by reasons why such measures may not be useful. Measures such as the Maudsley Obsessive-Compulsive Inventory (MOCI) and Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) featured in strong answers. Some candidates wrote descriptive essays which did not address the question set and could not be credited.

Question 10

Many candidates appeared to misunderstand the nature of the study by Knutson et al. (2007) on pre-cognitive decisions, on which this question was based. This study showed that the brain has made a decision before the person becomes cognitively aware of that decision. The common assumption is that the brain thinks one thing, yet cognitively a person might make a different decision. Appropriate answers often focused on the strengths of using fMRI to understand more about how the brain works (when making consumer decisions), whilst also arguing that when shopping it is not possible to take the fMRI along when the person would make a decision.

Question 11

Answers achieving the highest marks focused their attention on the arguments for and against the use of biochemical tests to measure adherence and threw into their discussions evidence from relevant research such as that by Roth and Caron (1978). Some candidates gave descriptive essays about adherence to medical advice followed with evaluation of their description, which did not address the question set. There are no AO1 marks for this question; all the marks are AO3 and so the question must be answered appropriately.

Question 12

The most logical approach to achieve a good answer was to provide a number of similarities between the theories of Maslow and Alderfer and then to consider all the differences (or vice versa). A few candidates took this approach and scored high marks. However, many candidates merely described Maslow's hierarchy and then described Alderfer's ERG theory without attempt to extract what was similar or different, and therefore did not score high marks.