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 their responses. The essay (final question) requires four evaluation points to be in depth (two strengths and two weaknesses) with at least one of these about the named issue. In depth tends to be having two examples of a particular concept or to support an evaluative point. Credit is limited if the named issue is omitted or just described.

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

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

Candidates 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. In addition, candidates need to know precise details about a procedure. This means presenting a level of detail about the procedure that would mean the study could be replicated. Generic 'stories' about a procedure will not gain credit.

Candidates should be able to give full definitions of terms listed in the syllabus and provide full assumptions for all four approaches.

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.

Section B was the strongest predictor of overall success for this examination paper with **Question 1** and **6** in ***Section A*** also correlating strongly with overall success.

General comments

The marks achieved by the candidates sitting this examination covered a wide spread of possible marks. However, three-quarters of the candidates scored 27 marks or less. 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 several blank responses in this series (every single item had blank responses). As positive marking is used, candidates should attempt all questions even if they are unsure of the response they are providing.

Many candidates could not differentiate between a result and a conclusion. A result is the collected data from a study that has been analysed via descriptive and inferential statistics. A conclusion is a generic commentary about what the results actually tell us, linked to the aim/purpose of the study.

Finally, there were several candidates who provided blank responses to any question related to a 'new' core study. It is essential that candidates have studied the correct syllabus – in this case the animal studies are Hassett and Fagen, not Yamamoto and Pepperberg.

Comments on specific questions

Question 1

(a) The majority of responses could correctly identify one of the cartoon characters. Common errors included naming other characters like SpongeBob Squarepants or Batman.

(b) Stronger responses could clearly identify two features of the sample. Popular choices included the number of participants, the age of the participants, and where they lived. Common errors included incorrect information about the sample used or listing features from a different core study (this was usually from the study by Bandura et al.). Candidates need to be familiar with the features of the sample from all 12 core studies. This had the joint second highest rate of blank responses.

(c) A large majority of responses presented a result rather than a conclusion. Common results were linked to identification rates of human and cartoon faces. These could not be awarded any credit. A small minority of responses could provide the generic conclusion based on the results presented by Pozzulo et al. There were many responses that 'concluded' that children are better at identifying cartoon faces over human faces. Conclusions need to be based around psychological principles and assumptions that are logical and meaningful. Knowing that children can identify better may be useful in other domains, but not for the study by Pozzulo et al. Candidates need to know the difference between results and a conclusion. A result is the collected data from a study that has been analysed via descriptive and inferential statistics. A conclusion is a generic commentary about what the results actually tell us linked to the aim/purpose of the study.

Question 2

(a) A significant majority of responses could provide a full aim of the study by Milgram. The most popular choice was focused on being obedient to an authority figure when asked to do something harmful/against a person's morals. There were some responses that simply stated 'to test obedience' but this could not be given credit as obedience was in the question. Candidates need to be wary about using words direct from the question in their response, as this cannot show understanding to an examiner.

(b) The majority of responses could identify a way in which the candidates were deceived in the study by Milgram. Popular choices included the use of a confederate, an incorrect aim or that they were led to believe that the person was receiving electric shocks. There was a substantial minority of responses that had clearly not studied the correct Milgram study. The core study on this syllabus has no vocalisations from the confederate – all of it was 'pounding on the wall'. It is essential that the correct Milgram study is being covered in centres.

Question 3

(a) A minority of responses could provide a result with a meaningful comparison. The most popular choice was comparing male and female monkeys. However, many responses focused on frequency or preference. The question was about duration of play. Therefore, these responses could not be awarded credit as they were not answering the question set.

(b) As with **Question 3(a)**, only a minority of responses could provide a result with a meaningful comparison. Stronger responses could present the correlational result based around plush toys and could be awarded maximum credit. However, there were many responses, like with **Question 3(a)**, where the candidate focused on frequency or preference (again this question was about duration of play). It is essential for candidates to know all key results from all 12 core studies.

(c) The majority of responses could identify a potential weakness of the study by Hassett et al. Popular choices included generalisability, mundane realism and categorisation of toys. Stronger responses would then provide a contextualised example from the study to explain why it was a weakness and these were awarded two marks. There were a range of weaknesses that were not creditworthy including arguments about cruelty, not being able to give consent, or not being able to generalise to humans (which was not the purpose of this study as they already had a comparison human child group).

Question 4

(a) Stronger responses could provide the definition for mallet aggression as used by Bandura, however, these were not common. Responses tended to be incorrect stating that it was the use of the mallet against the Bobo Doll. Candidates are required to know all behavioural categories used in this study.

(b) The majority of responses could identify one feature of Social Learning Theory. The most popular feature was observation. Many responses could also provide an example from the study about their chosen feature. However, most responses could not outline the feature chosen and many responses outlined many features of Social Learning Theory instead of focusing on one (as required by the question). To improve on this, candidates need to read the question carefully and ensure that their response is as focused as required by the question – in this example, **one** feature.

Question 5

The majority of responses provided examples of everyday life situations where doodling might help. Popular choices included in the classroom or at work. However, many responses presented an implicit 'how' in terms of stating 'allow doodling in the classroom or in the workplace during a meeting'. The term 'doodling' was in the question, so candidates needed to go beyond copying a word to show explicit understanding of what is involved in doodling to get above two marks. Stronger responses provided explicit information about how to doodle including providing paper materials, having shapes to shade etc. This showed explicit knowledge of what is involved in doodling so could be awarded up to five marks. It is really important for candidates to present explicit information to examiners to clearly show understanding of concepts already given in a question.

Question 6

The average mark awarded for this question was less than 2. Weaker responses tended to focus on all of the procedure rather than within the parameters of the question set. For example, describing what happened after the victim collapsed. Many responses provided a generic 'story' of what happened with no specific details about what the team of students were expected to do. It is essential for candidates to read questions of this type carefully to see from which two points their response should cover. Stronger responses could clearly describe the procedure from the two points highlighted in the question, providing a series of logical procedural points with specific detail, to be awarded maximum marks. This had the joint second highest rate of blank responses.

Question 7

The majority of responses could clearly identify two problems related to the use of children in research. Popular choices included potential ethical issues, lack of understanding and lack of linguistic skills. Stronger responses presented clear examples linked to the study by Baron-Cohen to suggest why they were potential problems. However, a large minority of responses did not provide contextualised arguments/examples based on the study by Baron-Cohen et al. Contextualised arguments/examples included not being able to understand the words used as the four choices in the eyes test and not having enough environmental/social experiences of emotions to be able to accurately understand/choose the correct emotion.

Question 8

There were a range of responses to this question. Stronger responses could clearly argue for either Jack or Daphne using a range of examples from the study. The more popular choice was to argue that Jack was correct about the study by Perry et al. being ethical. For responses that chose Jack, popular arguments were about the study clearly obtaining informed (valid) consent and that it was only a simulation so no actual personal space was being invaded. For responses that chose Daphne, a significant proportion tried to argue about deception. Double-blind studies use justifiable deception, so they are ethical (Daphne was arguing that the study was unethical) so the participant not knowing if they were being given oxytocin or a placebo is not deception in terms of ethical guidelines. For questions like this, candidates need to be able to provide correct examples and arguments.

Question 9

(a) A minority of responses could outline one dream reported in the study by Dement and Kleitman with a smaller proportion outlining the required two dreams. Popular correct choices included tomato throwing, climbing a ladder, and playing basketball. Many responses mixed together the reported dreams, for example, climbing a ladder the bottom of a cliff. There was a significant number of responses that created dreams not reported like flying, going to the dentist, or being late. It is important for candidates to know all key findings from all 12 core studies. This was the question with the highest proportion of blank responses on this paper.

(b) Stronger responses could clearly explain two differences. Popular choices included the compulsory sample of participants, brain measurement techniques, and experimental design. To improve responses to this type of question, candidates need to choose comparison points that can be developed and explained, using examples from both studies to explain each difference. For example, explaining the use of different brain measurement techniques with a focus on the function of each technique with clear examples from both studies. However, stating that each study had a different aim does not allow the response to be detailed so will always only achieve Level 1, neither does that they had different ethical guidelines to follow if choosing Hassett et al. as the second study. Candidates need to choose carefully what the comparisons are ensuring that they are logical and can be explained fully, using examples from both studies. It is also very important to read the question to see what can or cannot be used on the response. In this case, the candidates were told to refer to the sample, yet a minority of candidates did not use the sample in their responses and were therefore awarded the Level of their best 'difference' only. There was also a sizeable minority of responses that compared Dement and Kleitman to studies not from the biological approach. This had the joint second highest rate of blank responses.

Question 10

The strongest responses evaluated the study by Fagen 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 ethics, generalisability, observations, reliability, 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 Fagen et al. to explicitly support their point. These answers tended to score Level 5 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 Fagen 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 not in context which meant it could not be give Level 5. 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 are still following a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). 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 what quantitative data was collected. Some responses did not cover the named issue. There were also a large number of factual errors about the study by Fagen et al. presented as facts by candidates including the use of punishment, how the total training time reduced validity when this was actually 'minimal' and how the behaviour of 'steady' masked the success for trunk-up and trunk-here as the latter two were measured via the full trunk wash. It is essential that candidates choose

evaluation points based on what actually happened in the study. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, avoiding real world application. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples (plural) from the study to show clear understanding. An evaluation that is in depth tends to have at least two explicit examples from the named study for every evaluative point made. These are the requirements for a Level 5 response. The average response was Level 1 for this cohort.

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 in depth (two strengths and two weaknesses) with at least one of these about the named issue. In depth tends to be having two examples of a particular concept or to support an evaluative point. Credit is limited if the named issue is omitted or just described.

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

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

Candidates 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. In addition, candidates need to know precise details about a procedure. This means presenting a level of detail about the procedure that would mean the study could be replicated. Generic 'stories' about a procedure will not gain credit.

Candidates should be able to give full definitions of terms listed in the syllabus and provide full assumptions for all four approaches.

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.

Section B was the strongest predictor of overall success for this examination paper with **Questions 1**, **3** and **6** in ***Section A*** also correlating strongly with overall success.

General comments

The marks achieved by the candidates sitting this examination covered a wide spread of possible marks. However, half of the candidates scored 30 marks or less. 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 several blank responses in this series (fifteen items had some level of blank responses). As positive marking is used, candidates should attempt all questions even if they are unsure of the response they are providing.

Some candidates could not differentiate between a result and a conclusion. A result is the collected data from a study that has been analysed via descriptive and inferential statistics. A conclusion is a generic commentary about what the results actually tell us linked to the aim/purpose of the study.

Finally, there were several candidates who provided blank responses to any question related to a 'new' core study. It is essential that candidates have studied the correct syllabus – in this case the animal studies are Hassett and Fagen, not Yamamoto and Pepperberg.

Comments on specific questions

Question 1

(a) A slight minority of responses were able to present one of the other scales used in the study by Bandura et al. The most popular choice was physical aggression. Common errors included naming a scale in the final part of the study rather than the first stage, for example, mallet aggression. This had the highest rate of blank responses.

(b) Stronger responses could clearly outline how the children were assigned to one of the conditions in the study by Bandura. However, a significant majority of responses could not relate their response to the stem about the scales used by the experimenter and nursery teacher. These responses tended to focus on how many were in each condition which is just a description of the set-up rather than how they assigned them to the three main conditions. It is essential for candidates to know how researchers set up their studies.

(c) The majority of responses could identify one feature of the final experimental room. Popular choices included Bobo doll, a one-way mirror and naming a toy. Common errors included naming objects in rooms at different stages of the study that were not in the final experimental room. Some responses provided two examples of one type of toy. Candidates are required to know how a procedure is set-up for all 12 core studies. This had the third highest rate of blank responses.

Question 2

(a) A significant majority of responses could correctly identify the experimental design used. Some of these responses then went on to provide a reason why or an example from the study to complete the outline and be awarded two marks. Common errors included naming an incorrect experimental design or stating that the study was a laboratory experiment. Neither of these could be awarded any credit.

(b) A minority of responses presented a result rather than a conclusion. Common results were linked to recall rates using data to back up the claim. These could not be awarded any credit. A majority of responses could provide a brief, generic conclusion based on the results presented by Andrade, mainly about doodling affecting concentration. A minority of responses provided a full aim explaining what doodling may do to concentration or what cognitive mechanisms may have influenced recall. Candidates need to know the difference between results and a conclusion. A result is the collected data from a study that has been analysed via descriptive and inferential statistics. A conclusion is a generic commentary about what the results actually tell us linked to the aim/purpose of the study.

Question 3

(a) A slight majority of responses could provide a result. The most popular choice was the number of dreams recalled in REM. However, many responses did not provide a meaningful comparison to be awarded the second available mark. For example, comparing dream recall rates in REM and NREM. There were numerous errors presented based on what candidates believed Dement and Kleitman measured which could not be awarded any credit, such as dreams only occurring in REM. To improve, candidates need to know the correct results for all 12 core studies ensuring they present findings that were published in the original journal paper.

(b) Only a minority of responses could provide a result with a meaningful comparison. Stronger responses could present a comparison between 15 minutes and 5 minutes. However, there were many responses, like with **Question 3(a)**, where the candidate focused on one side of the result without any meaningful comparison. To improve, candidates need to know how results are presented. For example, there were many examples of a candidate stating that 78 per cent of participants could estimate, when the 78 per cent refers to the number of trials. It is essential for candidates to present full results with explicit comparisons with the correct parameters.

(c) The majority of responses could identify a potential weakness of the study by Dement and Kleitman. Popular choices included generalisability, mundane realism and ethics. Stronger responses would then provide a contextualised example from the study to explain why it was a weakness and these were awarded two marks.

Question 4

(a) Stronger responses could provide a full definition of obedience. A large majority of candidates cold at least present one of the two features of obedience – to follow an order plus an authority figure. A minority of candidates proposed that obedience is conformity and were not awarded any credit. Candidates need to know full definitions of all key terms highlighted in the 'psychology being investigated' section of the syllabus for all 12 core studies.

(b) The majority of responses could identify one ethical guideline and then provide an example from Milgram about how that guideline had been broken. Popular choices included deception, the right to withdraw, and informed consent. A minority of responses were awarded maximum marks as a second example of how the guideline had been broken was presented or an outline of the broken ethical guideline was given. Overall, there were many strong responses to this question.

Question 5

The majority of responses provided examples of everyday life situations where doodling might help. Popular choices included at work or in a therapeutic setting. However, many responses presented an implicit 'how' in terms of stating 'allow mindfulness to happen in the workplace or get patients to practice mindfulness'. The term 'mindfulness was in the question, so candidates needed to go beyond copying a word to show explicit understanding of what is involved in mindfulness to get above two marks. Stronger responses provided explicit information about mindfulness including meditation, yoga and body scanning. This showed explicit knowledge of what is involved in mindfulness so could be awarded up to five marks. It is really important for candidates to present explicit information to examiners to clearly show understanding of concepts already given in a question. The average score for this question was less than two out of five.

Question 6

The average mark awarded for this question was less than 2. Weaker responses tended to focus on different aspects of the procedure rather than within the parameters of the question set. For example, describing the procedure of choosing a target. Many responses provided a generic 'story' of what happened with no specific details about the actual line-up presentation. It is essential for candidates to read questions of this type carefully to see which part of a procedure is being asked for. Stronger responses could clearly describe the line-up presentation as covered in the original journal paper, providing a series of logical procedural points with specific detail, to be awarded maximum marks. This had the second highest rate of blank responses.

Question 7

The majority of responses could clearly identify two problems related to the use of children in research. Popular choices included potential ethical issues, lack of understanding and 'everything is a game'. Stronger responses presented clear examples linked to the study by Perry et al. to suggest why they were potential problems. However, a large minority of responses did not provide contextualised arguments/examples based on the study by Perry et al. Contextualised arguments/examples included the figures approaching being seen as a game, potential issues of ingesting oxytocin and with so many trials children would easily get bored.

Question 8

There were a range of responses to this question. Stronger responses could clearly argue for either Sabtu or Joyah using a range of examples from the study. The more popular choice was to argue that Sabtu was

correct about the study by Piliavin et al. being valid. For responses that chose Sabtu, popular arguments were about the study being a real-life setting and that no participant knew that the victims were confederates. For responses that chose Joyah, popular arguments included difficulty in controlling extraneous variables and a lack of diverse models. For questions like this, candidates need to be able to provide correct examples and arguments.

Question 9

(a) A minority of responses could describe more than one aspect of the psychology being investigated in the study by Fagen et al. Popular choices included operant conditioning and the use of rewards. However, a large majority of responses focused entirely on what Fagen et al. did in their study, rather than providing a generic account of the psychology being investigated. To improve, candidates need to know the psychology being investigated listed beneath all 12 core studies in the syllabus generically. There is only one mark available in these types of questions for an explicit example from the core study named. There were also a range of factual errors linked to this core study, for example, that classical conditioning was being tested, which is incorrect.

(b) Stronger responses could clearly explain one similarity and one difference. Popular choices to compare the studies on included quantitative data, use of rewards, generalisability and observational techniques. To improve responses to this type of question, candidates need to choose comparison points that can be developed and explained, using examples from both studies to explain the similarity and/or difference. For example, explaining the quantitative data of both studies would involve stating that direct comparisons can happen in both studies with examples of these comparisons from both studies. However, stating that each study had a different aim does not allow the response to be detailed so will always only achieve Level 1. Candidates need to choose carefully what the comparisons are ensuring that they are logical and can be explained fully, using examples from both studies. There were several examples where candidates did not choose a study from the learning approach which could only be awarded Level 0. To improve, candidates need to know how the 12 core studies are divided up into the four approaches.

Question 10

The strongest responses evaluated the study by Baron-Cohen 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 ethics. Common choices included ethics, generalisability, reliability, improvements from the original study, and mundane realism. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Baron-Cohen et al. to explicitly support their point. These answers tended to score Level 5 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 Baron-Cohen 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 not in context which meant it could not be give Level 5. 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 are still following a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). 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 ethics in an evaluative sense but did not fully explain why it could be a strength and/or a weakness, or simply described a range of ethical guidelines where only some were relevant to the study by Baron-Cohen et al. For example, the right to withdraw and debriefing are not mentioned explicitly in the study so candidates cannot assume how this was achieved. Some 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, avoiding real world application. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples (plural) from the study to show clear and explicit understanding. An evaluation that is in depth tends to have at least two explicit examples from the named study for every evaluative point made. These are the requirements for a Level 5 response. The average response was Level 2 for this cohort.

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 their responses. The essay (final question) requires four evaluation points to be in depth (two strengths and two weaknesses) with at least one of these about the named issue. In depth tends to be having two examples of a particular concept or to support an evaluative point. Credit is limited if the named issue is omitted or just described.

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

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

Candidates 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. In addition, candidates need to know precise details about a procedure. This means presenting a level of detail about the procedure that would mean the study could be replicated. Generic 'stories' about a procedure will not gain credit.

Candidates should be able to give full definitions of terms listed in the syllabus and provide full assumptions for all four approaches.

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.

Section B was the strongest predictor of overall success for this examination paper with **Questions 1**, **3** and **8** in ***Section A*** also correlating strongly with overall success.

General Comments

The marks achieved by the candidates sitting this examination covered a wide spread of possible marks. However, three-quarters of the candidates scored 21 marks or less. 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 several blank responses in this series (every single item had blank responses). As positive marking is used, candidates should attempt all questions even if they are unsure of the response they are providing.

Finally, there were several candidates who provided blank responses to any question related to a 'new' core study. It is essential that candidates have studied the correct syllabus – in this case the animal studies are Hassett and Fagen, not Yamamoto and Pepperberg.

Comments on specific questions

Question 1

(a) The minority of responses could correctly identify the sample size (either of the total or the size post-empathy split). Common errors included stating the sample size for a different core study. This had the joint second highest rate of blank responses.

(b) Stronger responses could clearly outline the term. Popular choices focused on the invisible boundary and choosing who can enter it. Some responses provided an example from the study by Perry which was awarded credit. There were a minority of tautological responses where the candidate wrote that person space was 'personal space'. These responses cannot be awarded any credit.

(c) A large majority of responses were not awarded credit. Many responses focused on what empathy is, how different participants reacted to each scenario or made up a way to measure empathy. To improve on this, candidates need to know the mechanisms used in all 12 core studies to define concepts, split participants into different groups etc.

Question 2

(a) A significant majority of responses could provide a full aim of the study by Bandura et al. The most popular choice was focused on whether children will imitate aggressive behaviour from an aggressive model. Some responses reported a finding from the study so could not be given any credit. A minority of responses provided information about the procedure of the study. It is important for candidates to know the difference between an aim and a finding.

(b) Stronger responses could provide the definition for aggressive gun play as used by Bandura et al., however, these were not common. Responses tended to be incorrect stating that it was the use of the gun to hit the Bobo doll or a behaviour using the gun that was not part of Bandura's definition. It is essential for candidates to know how behaviours were defined in this study. This had the joint second highest rate of blank responses.

Question 3

(a) A minority of responses could provide a result with a meaningful comparison. The most popular choice was that there was more same race helping when the victim was drunk for both races analysed in the study. However, many responses focused on different results about the 'drunk' victim without anything about the race of helper. These were not answering the question so could not be awarded any credit.

(b) As with **Question 3(a)**, only a minority of responses could provide a result with a meaningful comparison. Stronger responses could present a correct result with the most popular being the more males in the critical areas, the faster the time to provide help to the victim. However, there were many responses, like with **Question 3(a)**, where the candidate focused on a different result about 'victim helping' or stating that the more males in the critical area, the slower the response time for help. It is essential for candidates to know all key results for all 12 core studies.

(c) A slight majority of responses could identify a potential weakness of the study by Piliavin et al. Popular choices included generalisability, ethics and controlling extraneous variables. Stronger responses would then provide a contextualised example from the study to explain why it was a weakness and these were awarded two marks. Some candidates mixed reliability and validity and could only be awarded the identification mark.

Question 4

(a) The majority of responses could outline one feature of 'attention' with the most popular choice being about the ability to focus. Some responses could then continue to describe the term 'attention' in more detail sometimes with the use of examples from the study by Andrade. To improve, candidates should look at the mark tariff for the question and match the number of sentences to that. There were many examples of candidates writing two points about attention and nothing else which can only score a maximum of two of the three available marks.

(b) A minority of responses could identify just one feature of the sample. Popular choices included the age range and that they had just completed a previous study. A common error was identifying the sample as 'volunteer'. It isn't. There were a number of responses that re-named the two features in the question or described the two conditions used in the study. Neither of these could be awarded credit. To improve, candidates need to know all of the features of any sample used in any core study and then read the question carefully to ensure they are not presenting information already provided in the question.

Question 5

A minority of responses provided examples of everyday life situations where toy preferences might help. Popular choices included in the manufacture of toys or using certain toys with captive animals in zoos for rehabilitation. However, many responses presented an implicit 'how' in terms of stating 'making toys'. The term 'toys' was in the question, so candidates needed to go beyond copying a word to show explicit understanding of what is involved in 'toys' to get above two marks. Stronger responses provided explicit information about how to manufacture more wheeled educational toys for boys or combining wheeled and plush toys for girls. This showed explicit knowledge of what is involved in 'toys,' so could be awarded up to five marks. It is really important for candidates to present explicit information to examiners to clearly show understanding of concepts already given in a question.

Question 6

The average mark awarded for this question was less than 1. Weaker responses tended to focus on all of the procedure rather than within the parameters of the question set. Many responses provided a generic 'story' of what happened with no specific details about what techniques were used in the study by Fagen et al. to teach the elephants 'trunk up'. It is essential for candidates to read questions of this type carefully to see which points their response should cover. Stronger responses could clearly describe how the elephants were specifically taught, providing a series of logical procedural points with specific detail, to be awarded maximum marks.

Question 7

The majority of responses could clearly identify two problems related to the use of children in research. Popular choices included potential ethical issues, issues surrounding brain scanning and the confounding variable of grey matter development in childhood. Stronger responses presented clear examples linked to the study by Hölzel et al. to suggest why they were potential problems. However, a large minority of responses did not provide contextualised arguments/examples based on the study by Hölzel et al. Contextualised arguments/examples included not being able to keep still during an MRI scan and treating the mindfulness program as a game.

Question 8

The average mark awarded for this question was less than 1. There were a range of responses to this question. Stronger responses could clearly argue for either Luis or Ava using a range of examples/applications based on the study. The more popular choice was to argue that Luis was correct about the study by Milgram having application to everyday life. For responses that chose Luis, popular arguments were about the study, including how real-life situations could be explained using parts of the study. For responses that chose Ava, many focused on the controlled nature of the task, and that it is very unlikely to occur in everyday life. There were many answers that gave an anecdotal 'story' approach to the answer providing everyday examples of obedience which could only be awarded minimal credit. For questions like this, candidates need to be able to provide correct examples and arguments.

Question 9

(a) A minority of responses could describe one feature of either group mentioned in the question. The most popular choice was to describe Group 1, and within that, popular features included being diagnosed with AS/HFA, the number of participants and the fact they were all male. It was rare to see Group 2 being described. However, there were many responses that mixed up Group 1 with a different core study or provided information about what the sample 'did' in the study resulting in no credit being awarded.

(b) Stronger responses could clearly explain two similarities. Popular choices included the experimental design, quantitative data and lack of mundane realism. To improve responses to this type of question, candidates need to choose comparison points that can be developed and explained, using examples from both studies to explain each similarity. For example, explaining the quantitative data of both studies would involve stating that direct comparisons can happen in both studies with examples of these comparisons from both studies. Candidates need to choose carefully what the comparisons are ensuring that they are logical and can be explained fully, using examples from both studies. It is also very important to read the question to see what can or cannot be used in the response. In this case, the candidates were told not to refer to the sample, yet a minority of candidates did use the sample in their responses and were therefore awarded the Level of their best 'similarity' only. There was also a sizeable minority of responses that compared Dement and Kleitman to studies not from the cognitive approach. It is important for candidates to know which approach each of the 12 core studies 'belongs' to. This question had the highest rate of blank responses.

Question 10

The strongest responses evaluated the study by Saavedra and Silverman in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of quantitative data. Common choices included ethics, generalisability, reliability, 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 Saavedra and Silverman to explicitly support their point. These answers tended to score Level 5 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 Saavedra and Silverman 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 not in context, which meant it could not be give Level 5. 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 are still following a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). 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 what quantitative data was collected. Some responses did not cover the named issue. It is essential that candidates choose evaluation points based on what actually happened in the study. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, avoiding real world application. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples (plural) from the study to show clear understanding. An evaluation that is in depth tends to have at least two explicit examples from the named study for every evaluative point made. These are the requirements for a Level 5 response. The average response was Level 1 for this cohort.

PSYCHOLOGY

Paper 9990/21
Research Methods

Key messages

- Candidates frequently scored lower marks on questions where the stem used the command word 'explain', which requires more than just a statement. For example, many lost marks for not giving a detailed enough answer to a 2 or 3 mark 'explain' question but gave two basic suggestions instead.
- Candidates need to look for 'in this study' or other question prompts that indicate the need to contextualise their answer. Without this elaboration, marks will be limited.
- It is important that candidates acquire an understanding of the basic principles of effective tabulation, such as the correct use of headings and units.
- Many candidates did not understand how case studies are conducted and included details of an IV/DV and other participants. Candidates need to be prepared for this question, with a clear understanding of the four required features for each of the methods on the syllabus.

General comments

There were some topics that candidates seemed to know well on this paper, including recall of the sampling method used by Milgram, an understanding of quantitative data, the calculation of the range and identifying an independent and dependent variable. Many candidates seemed to find a number of aspects of research methods challenging, with many scripts unable to access high marks.

Comments on specific questions

Section A

Question 1

(a) This question part was well done, with most candidates being able to identify the sampling technique. A common error was to suggest opportunity sampling.

(b) This was another well answered question part with the majority of candidates scoring 2 marks. Some candidates suggested that flyers were put up around the university so did not gain the definitive mark for 'an advisement'/'direct mail'.

Question 2

Candidates found this question challenging, commonly giving the guidelines of protection from harm or number as opposed to the required species. Candidates who gave the right answer often needed to add relevant detail to earn the second mark. The most common correct response was identifying that researchers should not use endangered species, but this did not allow candidates to expand on their answer.

Question 3

(a) The majority of candidates were able to state 'quantitative' and, more often than not, follow through with 'numerical'. Some candidates were confused and stated 'qualitative' but went on to detail 'number' as the justification. A common error was to repeat the question by justifying with 'counting words' which could not earn credit.

(b) Many candidates began their explanation but did not achieve the second mark as their response was often generic, i.e. not linked to the scenario. Candidates need to read the question carefully, looking for 'in this study' or other question prompts asking for contextualisation of the answer.

Question 4

(a) Many candidates were able to achieve 2 marks here. A significant minority of candidates did not know what an independent groups design meant and attempted to explain that this was when participants work alone and not in a group. Another error was to correctly describe a repeated measures design, which was incorrect.

(b) Performance on this question part was slightly lower than seen in **part (a)** and candidates struggled to give a creditworthy description, usually only gaining one mark for stating the actual control condition of silence/no background noise. Many candidates described controls used in experimental studies to combat extraneous variables rather than explaining a control condition.

(c) Candidates often repeated the stem here, or repeated the procedure, which were not creditworthy responses. Stronger answers gained marks for stating observation and/or timing how long it took them to complete the tower. When responses did not gain full marks, this was typically due to a lack of detail.

Question 5

(a) Many candidates did not fulfil the 'both' requirement of this question. They did not address the question as to why it was necessary for children on the one hand and adults on the other to consent. A common mistake by candidates who gained 2 marks but not 3 was to omit an explanation of what informed consent is.

(b) Many candidates misunderstood the 'explain one way' command as 'explain why' and gave reasons why it was important for foils to be similar to targets. Some candidates suggested that they need to be in the same cartoon style which gained no marks or gave vague answers such as candidates' 'similar looks and appearance'.

(c) Most candidates achieved some marks here. A common error was insufficient labelling of the table, e.g., not stating 'percentage of correct identification' or, if they did, using it twice as a column descriptor which was a repeat and not creditworthy. Many also included values outside the acceptable range of error for reading from the graph. A small number of candidates drew bar charts or wrote paragraphs, which were not creditworthy. A significant minority of candidates omitted this question.

Question 6

Many candidates found this question challenging. The descriptions of participant observation and non-participant observation were often tautological, and many confused this with covert and overt observations. The most common mistake was to suggest that these were when the researcher took part or did not take part in the study. Some candidates were able to effectively use core studies to support their description, and Hassett et al. and Bandura et al. were both used for non-participant observation. Responses were more limited in using supporting studies for participant observation, often confusing the role of the observers in Piliavin et al. or omitting to explain that they appeared to be passengers, or suggesting the teacher was the observer in Milgram. There were some novel attempts at descriptions, some of which worked well.

Question 7

(a) Candidates seemed to find this question challenging. 'Explain which…would be most appropriate' is asking for a justification of why one would be chosen over another – in this case why one measure of central tendency would be chosen rather than another in each case. Very few candidates were able to describe or differentiate between quantitative data that was on a scale (for the mean) or in categories (for the mode).

(b) (i) This question was answered well by the majority of candidates. A minority of candidates incorrectly described the procedure for calculating the mean average.

(ii) Where candidates achieved a mark, the most common response was to relate it to the spread of data around the mean. Few candidates went on to make a comparison with the range. A smaller number gained a mark for identifying that the mean takes all the scores in the data set into account. These candidates were more likely to gain the second mark (for a comparison to the range using only two scores). Some responses referred to the relevance of outliers.

(c) Repetition of the stem was frequent here, with little additional information to explain 'why'. For example, candidates often rewrote the hypothesis, or repeated 'direction'. Another common mistake was to refer to correlations.

Question 8

(a) (b) These question parts were answered well by the majority of candidates. Two common mistakes were to have the IV and DV the wrong way round or to state 'children learning to read' for **part (b)**, which was a repetition of part of the stem rather than the variable being measured.

(c) Many candidates misinterpreted this question, giving a general description of why the pictures and shapes were brightly coloured as opposed to relating it to research methodology. Candidates that did recognise the concept of this as a control, went on to use elements of the scenario well, often citing attention, to stop boredom and making a judgement in the effect this control would have had on the validity of the study. Stronger responses also included that this would ensure differences in the DV were due to the IV.

Question 9

(a) (b) The vast majority of candidates misinterpreted the requirements of these questions, simply repeating the stem, i.e., 'participant' or 'situation'.

For **part (a)** there was little understanding of what constituted a participant variable. Those that did relate it to females, did not justify it in terms of their uniqueness. Candidates achieving the mark usually related their answer to personal lives.

Again, for **part (b)**, repeating the stem was common, with candidates focusing on the factory but not explaining that the factory was the environment/setting which was causing the problem.

Candidates may benefit from learning key words to answer this type of question. For example, knowing that participant variables relate to internal factors or individual differences and situational variables to external or environmental factors. This would help with understanding and identifying these variables as well as defining or explaining them.

(c) (i) More candidates were able to gain some credit on this question. Responses typically only gained 1 mark here as they lacked detail. Common suggestions given were either making sure the factory promoted both genders equally, selecting only females with fewer at-home responsibilities or helping the female workers with their at-home responsibilities. To gain full credit, candidates needed to expand their answers to include how questionnaires could be used to gain the necessary information or that the strategy would ensure that any differences in satisfaction would be due to gender only.

(ii) Typically, candidates offered one of two ideas. One of these being the idea that male workers may be unhappy, though these responses often only gained 1 mark for lack of detail. Stronger answers explained how this could have an effect on satisfaction levels and therefore validity would be reduced. Alternatively, candidates referred to a lack of generalisability and were more likely to gain full marks. Where candidates did not gain credit, their suggestions tended not to relate to issues that were created for Heng as a researcher or for the study, i.e. they were not answering the question.

Section B

Question 10

(a) Candidates did not appear to have a secure understanding of the case study method. The most common problem here was candidates trying to apply an experiment, especially the inclusion of two conditions, such as using boys with ASD diagnoses and without. Whilst of course we can use parts of the experimental method within a case study, this is not a complete response.

Stronger responses included some good knowledge and understanding of how data could be collected and what information could be useful, and mentioned triangulation and how qualitative and quantitative data would be interpreted.

Common mistakes – apart from focusing on an experiment – were the lack of two or more research methods and the lack of background information about the case/how to gather it (such as talking to parents, school etc.). It is essential that candidates are prepared for this question and have a clear understanding of the four required features for each method they can be asked about. Tackling each one will ensure that marks are improved.

(b) (i) Candidates need to remember to justify their answers to achieve full marks. A significant minority of candidates misread the question as how they could have improved validity rather than how it is already valid. Other candidates gave answers linking to reliability instead of validity so gained no credit. Stronger answers opted for explaining controls used or the type of data collected.

(ii) Candidates who discussed inter-rater reliability in their response rarely gained full marks as they were simply using the term rather than explaining how records between observers should be compared and strategies employed to ensure they were similar. Candidates who accessed marks on this question typically gave answers relating to using a more standardised procedure.

PSYCHOLOGY

Paper 9990/22
Research Methods

Key messages

- Candidates need to ensure that their responses are focused on the questions within the exam paper. There was more than one instance where candidates had misread the question and provided responses which were not creditworthy.
- Candidates need to ensure that they understand the expectations for different command words used on the paper. For example, 'explain' questions often scored poorly due to the lack of justifications provided within responses.
- Similarly, questions which asked candidates to explain benefits/problems of a particular method often scored lower due to the tendency to 'describe' rather than explain.
- Candidate responses showed a number of knowledge gaps of studies, with responses often using the wrong study in a response. Candidates need to be reminded that any of the studies in the syllabus can be used within the examination to show their knowledge of research methodology.
- It is worth noting that candidate responses for the higher (6 and 10 mark) tariff questions showed good knowledge and understanding of correlations and observations. There were a number of thoughtful responses for the extended essay question, and candidates should be commended for their performance on these questions.

General comments

This was the first summer series for the new Psychology syllabus. Responses to this question paper provided the full range of marks, showing a high level of knowledge and understanding across many areas of the syllabus. Where performance was limited, it was due to a lack of knowledge of studies, or specific research methods terminology. This was clear when looking at questions on case studies (**Questions 1a** and **b**), hypotheses (**Questions 3a** and **3b**), and control tasks (**Questions 5a** and **5b**). Candidate responses also showed gaps in knowledge when referring to terms such as validity and reliability, with some mixing up their responses. Candidates showed their ability to structure extended responses (**Question 10a**) with many able to produce thoughtful procedures which incorporated all the elements required within the question. It is clear that centres have prepared candidates well for the exam. For future series, candidates need to ensure that they have a good understanding of command words, key research method terminology such as validity and reliability, and the studies which have been named in the syllabus.

Comments on specific questions

Section A

Question 1

(a) (i) This 2-mark question required responses to state two features of a case study. The answer to this question was definitive to the features named on the specification; single unit/detailed. Many responses were able to state at least one feature and therefore were able to access some marks. There were fewer responses which were able to name both features. The most common error was that many responses suggested that triangulation and longitudinal study were key features; these were not creditworthy.

(ii) This 2-mark question asked candidates to link their stated features to the Salvedra and Silverman study. Where candidate responses had achieved both marks for **Question 1ai**, they often also achieved both marks for this question. Most responses achieved at least 1 mark for being able to identify the 9-year-old boy used within the study, but then often did not achieve the second 'detail'

mark, due to just suggesting it was a 'detailed study on phobias' which was not enough, or repeating points made in the first question.

(b) This 3-mark question asked candidates to explain a way to improve validity in a case study. A full range of marks was seen for this question although there were more responses at the lower end of the mark range. Many candidates showed some understanding of validity and were able to identify a way such as standardisation or the use of different methodology but were unable to elaborate on these ways to enable them to access all 3 marks. There were some strong responses, most often about triangulation, which often produced 2 or 3 marks. The most common error beside lack of detail, was a misunderstanding between reliability and validity which reduced the number of marks available for the response. Candidates must ensure that they fully understand research method terminology such as reliability and validity.

Question 2

(a) (i) This was a very accessible question for most candidates. The majority were able to identify the sampling technique as volunteer, and a significant amount were able to then go on and explain why this is the case. The most common responses used the idea that participants respond to an advert/were people who are willing to take part. Where performance was limited, it was due to responses repeating the word 'volunteer/self-selected' in the explanation, such as 'volunteer sampling as it is where participants see an advert and volunteer themselves'. This would only achieve 1 mark.

(ii) This 2-mark question asked candidates to suggest one weakness of the sampling method used. For this question candidates needed to identify a weakness and then link this to the scenario. The vast majority of candidate responses were able to achieve at least 1 mark for identifying the weakness, usually by correctly suggesting that a lack of generalisability is a weakness of volunteer sampling. The most common error was that many responses were generic and therefore were only able to achieve a maximum of 1 mark. To achieve the second mark, responses needed to link back to the scenario using words such as shoppers/town/responded to his advert.

Question 3

(a) Almost all responses found this 1-mark question accessible and were able to achieve the mark available. The answer for this question, 'directional hypothesis' was definitive.

(b) This 2-mark question asked candidates to explain their choice for **Question 3a**. Performance on this question part was mixed. Candidates needed to identify a weakness and then link this to the scenario. Creditworthy responses for 2 marks included: it says which IV level (older children) was better (DV)'. Many candidates did not achieve marks on this question due to just repeating the scenario. For example, that it shows the direction of the results or repeats that 4-year-old children make more mistakes than 8-year-old children. This is just repeating the wording in the question, or just lifting off the scenario, which is not creditworthy.

Question 4

This 4-mark question asked candidates to outline two ethical guidelines used within the Hassett et al study. For each point, responses needed to name a relevant ethical guideline and then link this to the Hassett study. Creditworthy ethical guidelines included pain and distress, reward, number, species, and housing. Most responses could correctly identify two ethical guidelines for 2 marks, with the most common ones being housing and pain/distress. The most common error for this question was that responses did not elaborate on the guidelines or did not link them back to the study. In addition, some candidates, when referring to housing, would just provide a description of the housing for the monkeys rather than linking it to the ethical guidelines which meant that these responses could only achieve 1 mark for that guideline.

Question 5

(a) This 1-mark question asked candidates to identify the control task in Baron-Cohen. Very few candidates recognised that the control task was the gender recognition task, with the most common error being that responses suggested that the glossary was the control task.

(b) This 1-mark question asked candidates to outline what participants were required to do in the control task named in **Question 5a**. Inevitably candidate responses which were unable to identify

the gender recognition task were very rarely able to access this mark. Candidates who had answered **Question 5a** correctly, were mostly able to achieve the mark available. The most common error in responses was just repeating what they had said in **Question 5a** without making it clear that these were from photographs of eyes, often just saying 'to see if they were male or female' which was not enough for the mark.

Question 6

This 6-mark question required candidates to describe how variables in correlations could be measured. Awarded marks could come from specific methodologies such as observations/interviews, specifics from within these named methodologies such as use of behavioural categories and likert scales and examples such as the AQ test score from Baron-Cohen et al. Candidate responses to this question were mixed. At the lower end of the mark range, responses often focused on descriptions of *types* of correlations such as positive and negative, as well as how correlations could be *displayed* such as through a scattergraph; these responses tended to either be awarded no marks, or 1 mark for identifying the AQ and Eyes test as a measurement technique. At the high end of the mark range, thoughtful responses were seen which looked at a variety of methodologies, how these methodologies could be used, and a number of studies which had used these methodologies successfully. It is worth noting that examples do not have to be from named studies, and there were some creative responses which used a variety of examples; the most common of which was sales of ice cream and temperature; which was indeed creditworthy.

Question 7

(a) (b) These 1-mark questions asked candidates to identify the open and closed questions from a list of four. The vast majority of candidates gained both the marks for these questions and no real issues were found. The most common error was to choose 'write about whether you would like to be a doctor' as a closed question which is incorrect.

(c) (i) This 1-mark question asked candidates to state one strength of Question E (which was a closed question). Creditworthy responses included objectivity, the ability to statistically analyse and allows for comparisons between participants. The vast majority of candidate responses were able to access the mark available for this question. The most common error was to use terminology such as 'quantitative data' without actually saying why this would be a strength.

(ii) This 1-mark question asked candidates to state one strength of Question F (which was an open question). A creditworthy, and by far the most common, response here was the idea of in-depth/detailed data. This question performed better than **Question 7ci** above, with the vast majority of candidate responses accessing the mark available. The most common error, similar to **Question 7ci** above, was for candidates to use terms such as 'qualitative data' without saying why this would be a strength.

(d) This 2-mark question asked candidates to suggest two questions that could be asked to investigate subject choices. This could be one open and one closed question, or two closed or open questions. Many candidates could achieve at least 1 mark on this question, usually through an appropriate open question. Where candidate responses did not achieve both marks it was, in the main, due to a lack of choices given for closed questions, such as likert scales or yes/no choices. A minority also produced questions which had no relation to the topic area stated in the question, such as asking for the age of the participant. It is worth noting that closed questions have to have choices, and open question have to generate genuine qualitative data through the use of words such as explain, describe, tell me about etc.

Question 8

(a) This 2-mark question asked candidates to explain one reason why Daku's conclusion may be incorrect. Responses needed to relate to interpretation of the data given on the bar chart. Reasons could include the small difference between girls and boys or the fact that obedience to older people was the same. The detail mark could then come from candidate responses suggesting that this difference was not significant or could be due to chance. Candidates seemed to find this question difficult. Although many responses could gain a mark for suggesting that obedience to older pupils was the same, candidates then struggled to extend their responses to achieve the second mark. Responses which achieved both marks often suggested that the difference was too small to make that sort of conclusion, as this difference could be due to situational factors not considered, which was appropriate for 2 marks.

(b) This 2-mark question asked candidates to explain one benefit of increasing the sample size in this study. For this question, candidates needed to give a reason such as to increase generalisability, and then give more detail of why this would be a benefit, such as increasing the variety of children with different obedience levels. Most responses were awarded at least 1 mark for this question for suggesting generalisability as a benefit. The most common error was not giving enough detail to achieve the second mark. Often responses would talk about generalisability in terms of getting a 'bigger sample'. Having a bigger sample does not guarantee that their results are more generalisable therefore there needs to be something in the response that talks about the diversity of the sample in terms of obedience levels/culture etc. which is relevant for the study.

(c) (i) This 1-mark question asked candidates to explain whether the problem stated in the scenario was caused by a situational or participant variable. Many responses did not achieve the mark on this question as they did not *explain* the answer and only stated it. For example, saying situational variable alone was not enough to get the mark, responses needed to say that this was because it was due to environmental/external factors.

The most common errors, besides just stating situational variable, were that responses would say 'it was due to the situation not the person' which is not creditworthy as it is tautological, trying a 'reverse argument' such as 'because it has nothing to do with the participant', and also repeating what was in the scenario such as saying 'because only boys waited longer not girls' which was not creditworthy. Candidates need to be aware of the different requirements for command words such as *explain* and *state.*

(ii) This 1-mark question asked candidates to say how the problem noted in **Question 8ci** would affect results. Despite the struggles candidates faced with **Question 8ci,** this was answered well with some thoughtful responses. Most were able to suggest that their disobedience would not be caused by the independent variable that the researcher was manipulating (gender) but levels of hunger/frustration/impatience/anger. The most common error for this question was writing a response which was vague and did not really answer the question. For example, some responses would just say 'because if boys waited the same time the results would change'. In itself, this is correct, the results may change but responses needed to say *why* this would happen, such as they would not become impatient etc.

Question 9

(a) This 2-mark question asked candidates to suggest one problem with Carol's measure of music and then give some detail about why this would be a problem. Many responses were able to access at least 1 mark for this question and indeed most of these would then go on to give some detail to get the second. Creditworthy responses included missing some of the beats due to it being loud (therefore lowering the reliability/validity of the measured) and not including other aspects of the music that may affect levels of exercise (such as the lyrics/genre). The most common error was suggesting that music could have slower and faster beats within one song. This is not creditworthy as she counted number of beats per minute therefore this would have been considered (so music with both types would have less beats per minute than a purely fast song).

(b) This 4-mark question asked candidates to explain two reasons why Carol obtained permission to gather her data, in relation to the ethical guidelines. To achieve the 4 marks, responses needed, for each point, to give a reason related to ethics and then link this reason to the scenario. Ethical guidelines which were creditworthy included privacy, harm, right to withdraw and informed consent but not confidentiality and deception. Most responses were able to access some of the marks available for this question, with many able to access all 4. The most common error was not linking the ethical issue stated to the scenario which meant that the maximum a response could be awarded was 2 marks. It is worth nothing that the use of words such as 'observation' and 'recorded' were not enough of a link to the scenario but such words/phrases such as (sports club) members, and exercising were enough.

(c) This 2-mark question asked candidates to explain one reason why it is important that Carol visits the sports centre at different times of day. To achieve the 2 marks, candidate responses needed to give a relevant reason such as different levels of fitness/different types of music/different types of people at different times of day, and then provide some linked detail related to their stated reason, such as people in the morning would therefore exercise longer no matter the type of music that

they hear. Candidate performance on this question was mixed. The most common error was the lack of detail given which meant that many candidates were only able to achieve the 1 mark.

Question 10

(a) This 10-mark essay question asked candidates to describe how Chloë could conduct an observational study using participant observation to record the variety of behaviours during meals at work. Responses were awarded the full range of marks, although there were more at the lower end.

Many candidates had a sound understanding of observations and were able to make relevant decisions about whether it should be covert or overt, structured or unstructured. There was a lack of understanding about what was meant by a naturalistic observation and this often meant that responses were not able to achieve the highest mark band.

Responses within the lower mark range were often able to give a list of behavioural categories which they were going to observe and were able to produce a basic procedure that Chloë could follow. At this level, some responses had significant gaps within the procedure which meant that it would not be replicable. Some responses also mistakenly suggested that they would sit apart from the participants or observe through cameras. This would not be creditworthy as the question required the research to be participant observation. In addition, at this mark range some responses suggested the use of questionnaires alongside observation which is also incorrect.

Responses within the higher mark ranges successfully described a procedure that would be replicable by other researchers. Most candidates at this level suggested a covert observation, and then described how this would be achieved. For example, they would suggest that the researcher should pose as a worker and sit with the participants and try and engage them in conversation whilst recording them on their phones. The behavioural categories described at this mark range were thorough and, for the most part, fully operationalised. Most responses at the higher mark range showed understanding of what is meant by a naturalistic observation although it was still slightly confused at times. Candidates' understanding of participant observation was clear and explicit.

It is essential that candidates are prepared for this question and have a clear understanding of the four required features for each method they can be asked about. This will ensure that in future series candidates are able to achieve the marks at the highest levels.

(b) (i) This 2-mark question asked candidates to describe one practical/methodological strength of the procedure. To achieve the 2 marks, candidates needed to identify a strength and provide some detail about that strength. Common strengths included the use of covert observation reducing demand characteristics and operationalisation of categories which would increase the replicability/reliability of the research. Performance on this question was, in the main, good. Most responses were achieved at least 1 mark, with a significant proportion achieving both marks. The most common error, similar to previous questions, was that responses did not give enough detail to achieve the second mark.

(ii) This 2-mark question asked candidates to describe one practical/methodological weakness of their procedure. To achieve the 2 marks, responses needed to identify a weakness and provide some detail about that weakness. Common weaknesses included inability to control extraneous variables and the use of overt observation increasing demand characteristics. As with the previous question, most responses were able to achieve at least 1 mark, with a significant proportion of these achieving both marks. The most common error was the lack of detail given which meant that access to the second mark for those responses was limited. In addition, some responses used weaknesses which were not linked to the procedure they had written. For this, and the previous question, it is important that candidates only use strengths and weaknesses that are explicitly linked to something they have already written in **Question 10a**.

PSYCHOLOGY

Paper 9990/23
Research Methods

Key messages

- Both the research methodology and core studies elements of the syllabus are important for this paper. Although entitled 'Research Methods', candidates are expected to be able to draw on their knowledge of the core studies to answer some of the questions.
- Some questions only require a simple definition or description. The majority, however, also include a requirement for detail or a link, e.g., to a core study or part of a scenario. Such extensions to answers were commonly omitted.
- There are some basic concepts that candidates clearly understood, but some others could have been better understood or expressed. These included concepts such as experimental and control conditions and correlations as well as matched pairs designs.
- With regard to the 'design a study' question, candidates need to be prepared and have a clear understanding of the four required features for each method on the syllabus.

General comments

Many candidates seemed unable to recall knowledge of the core studies or a number of aspects of research methods. There were some topics that candidates seemed to know well, including *repeated measures* and *operant conditioning*. There was a considerable number of blank responses and scripts seen. 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

Section A

Question 1

Most candidates could define a covert observation although few gave a clear example, with most just naming the Piliavin et al. study.

Question 2

The vast majority of candidates simply repeated the information from the question or gave conclusions from the study. This suggests that they had not understood the instruction to 'Outline what these two results *show*.' For the core studies, candidates need to know how the results can be interpreted.

Question 3

Candidates gave a range of answers here, with some identifying key features such as the deception or the prods.

Question 4

This question was generally well answered, with right-handed/left-handed and male/female being the most common answers. Often responses were too brief for full marks. For example, many candidates were able to identify that the sample was too small to be representative but did not extend this to explain why that was so i.e. to say that this would mean the sample would lack variability/diversity.

Question 5

(a) There were some clear definitions of objectivity and a wide range of appropriate examples. Some candidates, although able to explain objectivity, were only able to do so generically. A minority of candidates did not earn credit for merely describing objectivity as the opposite of subjectivity.

(b) There were some occasional strong answers to this question part, for example identifying that the researchers knew what the correct answer was. However, the majority of responses just identified features, such as saying '*cartoon characters*' with no link to the question at all.

Question 6

Candidates appeared to find this question difficult. In particular, matched pairs was not well understood by candidates. Many successfully identified IQ as an example of matching from the Baron-Cohen et al. study. A common error was to say that candidates were matched and put into the same level of the IV, rather than into different levels. Repeated measures was much better understood with the majority of candidates able to state that 'repeated measures is where the participants take part in all levels of the IV.' Weaker responses simply stated that 'repeated measures was doing a study again', so did not earn credit. As with **Question 1**, studies were often named as examples without any relevant information being given.

Question 7

(a)(b)(i) These two question parts were not well answered. Candidates appeared to find difficulty with expressing the difference between an experimental condition and a control condition. For example, candidates typically made errors such as describing the experimental condition as the 'one that changes' and the control condition as the 'one that does not change'. Some candidates misunderstood the question and identified what the experimental condition would be in this study, or gave definitions of a control, such as 'a thing that is kept the same'. Some candidates did understand the distinction between the two and explained both clearly. Candidates may benefit from learning key words to answer this type of question, such as that the 'experimental condition is a level of the IV that is manipulated'.

(ii) The majority of candidates correctly identified '*no music'* as the control condition here. There were a range of other responses that were not creditworthy, including the type of music, the volume of music or even the type of headphones.

(c) The majority of candidates identified the type of experiment as a '*field experiment'*, although often going on to say that this was because it was conducted in a natural environment. Some candidates incorrectly identified the study as a natural experiment, or as a correlation or case study. Candidates were less often able to correctly justify their choice to earn the second mark by linking their answer to Yashal's study as required by the question. Where they did so, this was most commonly by saying that they were at home.

Question 8

(a) This question part was well answered. Most candidates successfully made reference to reinforcement/reward, treat/food etc., as well as to luring the parrot into the cage, successive reinforcements and the gradual removing of reinforcements. Some candidates added irrelevant material, such as that rewards need to be associated with something such as a bell, so using a classical conditioning technique in combination with the operant technique. Some candidates also included irrelevant concepts from social learning theory in combination with rewards.

(b)(i) Whilst many candidates seemed to find this question difficult, there were some very good answers. Many candidates focused on individual characteristics of parrots such as their age or their learning abilities rather than the measurement. Candidates who did focus on measurement identified a range of appropriate issues. Where they did not, this was often because they described the procedure rather than responding to the question set.

(ii) Several candidates did not score marks here because they had not scored marks for the previous question part. Many answers only earned one mark as they needed to be developed further. For example, identifying days as too broad a measurement for the previous question and simply saying 'measure in hours' here. Some candidates were able to identify a successful way and give detail of how it could be solved for 2 marks.

(c) (i)(ii) The first part of this question was relatively well answered with candidates suggesting that the presence of other staff would make the parrot more comfortable or would provide access to greater reinforcement. Some interpreted the question slightly differently and explained that training like this would be better because this was a more typical environment for the parrot. The second part of the question was often reversed so rather than explaining why the parrot would learn better in the absence of other staff, many candidates explained why the parrot would learn worse in the presence of other staff (BOD credit was given for these). Many candidates were able to score full credit here, typically with responses suggesting additional people would distract the parrots and prevent them from focusing on the learning task.

Question 9

(a) This question was generally well answered with most candidates able to explain that artificially creating pollution in a laboratory would break the ethical guideline of protection from harm.

(b) This part of the question was less well answered with many candidates simply stating that it would be less harmful outside. Stronger answers showed understanding that the pollution was not being artificially created for the purposes of the study and that the participants would be in these environments regardless.

(c) A small number of candidates misunderstood this question and explained the effect of uncontrolled variables in generic terms. Most candidates understood what was required and made suggestions about prior experiences, pre-existing mood disorders, other things that happened whilst on the trip, weather and so on. To gain full marks, candidates needed to comment on how they would affect the measurement of emotion rather than simply identifying the variables.

Section B

Question 10

(a) Most candidates did not seem to fully understand what a correlation is. Many answers started by identifying groups, for example of shy and not shy people, with no indication of how this would be determined, or started by identifying shyness as an independent variable which would have an effect on happiness as a dependent variable. Very few responses met the criteria of continuous data for both variables, although many gave detailed descriptions of how to measure one of these variables. To improve marks on this question, it is critical that candidates are aware of the required features for each research method as stated on the syllabus. For correlations, the first of these is the 'two co-variables', yet almost no candidates defined the two variables given in the question.

(b) (i) Answers to this question part had a tendency to be generic. Responses often referred to comparisons between groups etc., which were inappropriate or to sampling or ethical issues which were excluded by the question.

(ii) Responses often lacked detail to this question, simply identifying a generic weakness. The most common responses were demand characteristics/social desirability/lying.

PSYCHOLOGY

Paper 9990/31

Specialist Options: Approaches, Issues and Debates

Key messages

Questions 1, 3(a), 5(a), 5(b), 7(a)(i), 7(a)(ii), 9(a), 9(b), 11(a), 13(a), 13(b), 15(a)(i) and 15(a)(ii)

These questions in this exam asked candidates to apply an area of the syllabus (theory, technique/treatment, disorder, etc.) to explain how it is relevant to a particular scenario or context. It is important that candidates are aware of the bullet points in the syllabus. It would be helpful for candidates to write revision notes with the title of the topic area and bullet point at the top so that they can identify which part of the syllabus these types of questions are referring to. Candidates should also refer directly to the scenario/context in the question in their response.

Questions 3(b), 7(b), 11(b) and 15(b)

These questions in this exam asked candidates to evaluate the suggestion, such as the technique/treatment, that was outlined in the candidate's response to **part (a)** of the question. In this exam, this type of question asked the candidate to evaluate the technique outlined in **part (a),** such as with a strength, weakness or a problem a psychologist could have when they investigate the scenario/context given in **part (a)**. It would be helpful to candidates when revising to learn strengths and weaknesses of the theories, techniques, self-reports, treatments, etc. they have learned and put these into their revision notes. They should also practice explaining the evaluation point in the context of the question.

Questions 2, 6, 10 and 14

Part (a)

These questions could ask the candidate to outline a theory, study, technique/treatment or self-report used by psychologists that is named in the syllabus or outline one of the issues and debates, possibly with an example from the syllabus content. The revision technique outlined previously in this report will aid candidates in learning the syllabus material.

Part (b)

It would also be useful for candidates to write revision notes where they define the issues/debates and prepare a strength and a weakness of each issue and debate to prepare for **part (b)** of this type of question. These questions in this exam were worth 2 marks for each part of the response and therefore a short response is appropriate.

Questions 4(a), 8(a), 12(a) and 16(a)

These questions in this exam came from one or two of the bullet points in the syllabus. This exam can either ask the candidate to outline a key study from the syllabus or two studies, theories, characteristics/explanations/treatments 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 could be useful for candidates to create revision notes with the title of each topic area and the description in the bullet point as the header. Alternatively, candidates could create a mind map and put this information in the centre.

Questions 4(b), 8(b), 12(b) and 16(b)

This question will always ask the candidate to evaluate the studies, theories, characteristics/explanations/treatments of disorders or techniques described in **part (a)** of the question. The response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. 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 studies, theories, characteristics/explanations/treatments of disorders or techniques described in the **part (a)** of the answer.

The candidate must also provide some form of analysis to access Level 3 and above. 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. The response needs to explain the comparison/strength/weakness or counter-argument with examples from **part (a)** of the question. It was common for responses to state that two theories, for example, were 'similar' or 'in contrast' for an issue without any explanation as to why they could be compared in this way. This is limited analysis. 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 4 and 5 descriptors, it would be best to structure the response by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure the answer covers this requirement of the question.

A small minority of candidates did not evaluate using the named issue. Quite a few of the answers were structured by study/theory/treatment rather than the issue which often led the response to be quite superficial and repetitive. A number of the responses successfully included analysis. Candidates should be aware this question is worth 10 marks and so they need to include an appropriate amount of information.

General comments

The marks achieved by candidates for this session of the 9990 specification achieved across the full range of the mark band. Stronger responses showed good knowledge, understanding, application and evaluation. Weaker responses showed limited knowledge and understanding with brief, superficial and sometimes anecdotal answers. These responses often had limited evaluation and application skills.

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

The questions on clinical were the more popular choice of option, followed by health.

Comments on specific questions

Clinical Psychology

Question 1

There were many good responses to this question which asked for an explanation of how Sarah meets the diagnostic criteria for obsessive-compulsive disorder (OCD). The best responses were able to give the specific symptoms from the ICD 11 and link these to Sarah. For example, obsessions linked to germs and compulsions linked to washing hands for 15 minutes/washing coat, shoes and clothing. Other common symptoms included significant impact on everyday life with examples of poor sleep and Sarah isolating herself at work. Finally, some responses explained how the compulsion of hand-washing took up more than an hour a day as Sarah washed for 15 minutes every hour.

Weaker responses frequently focused on one or two symptoms with the most common being Sarah's obsession with germs, followed by her compulsion to wash her hands. A less successful strategy was to list the symptoms at the beginning in the first sentence without linking these specifically to Sarah's thoughts and behaviours. In addition, weaker responses did not give the specific wording/terminology for the symptom. For example, missing compulsions and just saying Sarah repeated behaviours.

Question 2

(a) There were some strong responses to this question with the candidate outlining what is meant by determinism and how that is distinct to free will. Clear definitions for determinism stated that behaviour is caused by factors that are out of our control/determined by internal/external factors. Many responses were able to give a good outline of free will with candidates stating that it is where behaviour is due to the individual's choice. Weaker responses had elements of the debate but were vague or unclear e.g. 'something is already determined for us' which was not creditworthy. Overall, candidates were able to describe the free will side better than the determinism side.

(b) Candidates who provided a creditworthy definition of determinism in **part (a)** were then able to explain a strength of applying the determinism to the biochemical explanation for depressive disorder, although the responses were often very brief with the strength being briefly outlined. The most common strength given was treatments with some giving examples of SSRIs. A few responses outlined the strength that this takes away the blame from the patient for their depressive disorder and some extended this by explaining that this would lead to less distress for these patients. A significant number of responses did not address the question and described how the biochemical explanation for depressive disorder is deterministic rather than explaining a strength, which was not creditworthy. It was noted that there was some overwriting in some responses where a description of the biochemical explanation was given, which was also not creditworthy.

Question 3

(a) Responses to this question covered the full range of the mark scheme. Those that achieved 3 – 4 marks outlined the process of cognitive restructuring the therapist would use with Olivia with specific reference to Olivia's symptoms outlined in the question. Most responses explained that cognitive restructuring involves changing thoughts from irrational to more rational. Stronger responses gave examples such as restructuring her thoughts that 'no one likes me' to 'I am a likeable person and sometimes my family might be distracted or busy and just appear to not like me'. A few responses also suggested how the therapy could help Olivia to manage her manic episodes with examples about the irrational thought of thinking your life is 100 per cent perfect.

Most responses focused on the depressive episodes and so were limited to a maximum of 3 marks for this question.

Weaker responses frequently gave a general outline of cognitive restructuring or mentioned challenging irrational thoughts without any examples from Olivia. Some weak responses gave lists of what any therapist might do, including making the patient feel comfortable and/or setting homework with no indication of their understanding of cognitive restructuring. These responses were not creditworthy.

(b) There were some full mark responses to this question that could give a clear explanation of a strength of using cognitive restructuring therapy with Olivia. The most common strength given was that the therapy does not have side effects compared to drug treatments with an example given of one or two side effects. Another good strength that was often awarded full marks was explaining how cognitive restructuring can provide Olivia with skills so that she can use it in the future if she has mania/depressive episodes. Weaker responses that achieved 1 mark often just stated 'no side effects'. Responses that were not creditworthy stated that the strength of cognitive restructuring is that it will help her overcome her mania/depressive episodes without any explanation on how.

Question 4

(a) Responses varied for this question and covered the full range of the marks available. Level 3 responses often gave clear details of the behavioural explanation of fear-related disorders with good use of appropriate terminology. For example, identifying the rat as the neutral stimulus which is paired with the unconditioned stimulus of the iron bar banging which initially led to the unconditioned response of fear to the banging. Once a few pairings were done, the rat became the conditioned stimulus and this would produce a conditioned response of fear towards the rat (without any loud noises). For operant conditioning, the use of negative reinforcement where the person avoids the phobia was outlined in some responses. There were also a few good descriptions of Freud's theory. For example, explaining that phobias are repressed to protect the ego into our unconscious and then displaced, the link to the phallic stage and the Oedipus

complex. Many responses gave an outline of Freud's study of little Hans. Some focused on the development of the horse phobia but many responses gave very long descriptions of the study including details which were not relevant to the development of Hans's horse phobia (or the candidate did not link the details to the phobia).

Weaker responses often lacked or used incorrect terminology. Many Level 1 and sometimes Level 2 responses included descriptions that were not relevant to the question, such as outlining cognitive explanation for phobias such as the Chapman and DeLapp study on blood phobias which is about CBT. In addition, some responses outlined Freud's theory about the development of personality rather than phobias. Incorrect responses explained that phobias develop during the anal stage and gave the psychodynamic explanation of OCD rather than phobias.

(b) Similarly to **part (a),** there was a variety of responses to this question and the marks achieved were frequently between Level 1 and Level 3. Most responses included the named issue of longitudinal and there were many that included clear examples and some analysis. A few impressive responses were able to distinguish between the benefits of the little Hans study being carried out for longer than little Albert. These responses often discussed both the strengths and the weaknesses of longitudinal research. Stronger responses included fewer issues in some detail with effective analysis, often focusing on 3 – 4 evaluation issues. Other common issues included use of case studies, reductionism versus holism, generalisability, determinism versus free will, validity, and individual and situational explanations. In addition, ethical issues were common which was appropriate for both studies. Many were able to discuss the psychological harm caused to little Albert and some discussed the ethical strengths of the studies as parental consent was gained. There was frequently a misunderstanding of how ethics applies to children in psychological research with candidates stating that there was no consent or right to withdraw in spite of the parent being present throughout the study.

Weaker responses did not contextualise their response. While some of the evaluation points were valid, the lack of context prevents candidates from achieving a higher band. Candidates who structured their response by strengths and weaknesses of the studies were often repetitive, as both responses were case studies and had a weakness in generalisations. Some responses provided too many issues with no depth in explaining why. This type of response often identified an issue and then stated that it applied to either one or both of the studies before moving onto the next evaluation issue.

Consumer Psychology

Question 5

(a) There were a number of good responses to this question and many achieved full marks. Full mark responses were able to identify that high self-monitors were image oriented/soft-sell and were able to outline a clear advertisement strategy that focused on image. For example, a celebrity wearing the coat, the coat makes the wearer look trendy and other people are amazed at how good the coat looks on the person in the advertisement. Weaker responses did not outline a clear enough strategy that linked the advertisement strategy to image. These responses usually could outline that high self-monitors appeal to image oriented advertisements, but the strategy was too brief or not attempted. Some candidates confused **part (a)** and **part (b)** of this question and suggested that high self-monitors were quality focused which was incorrect and not creditworthy.

(b) Similarly to **part (a)** of this question, there were a number of good, full mark responses. These responses were able to identify that low self-monitors were quality oriented/hard-sell and were able to outline a clear advertisement strategy, featuring details of the quality coat. For example, that the coat is warm for the winter, the materials the coat is made out of are good quality or the coat is hard-wearing and will last the wearer a long time. Weaker responses often did not outline a clear enough strategy that linked the advertisement strategy to quality. These responses usually could outline that low self-monitors appeal to quality oriented advertisements, but the strategy was too brief or not attempted. Similar to **part (a)**, some candidates confused **part (a)** and **part (b)** of this question and suggested that low self-monitors were image focused which was incorrect and not creditworthy.

Question 6

(a) Many responses were able to give a clear outline of 'satisficing' as a model of consumer decision-making to achieve full marks for this question. Strong responses included being able to outline that satisficing is when you buy something that is just 'good enough' or meets your basic needs. Candidates who did well made the link between the concept and the consumer decision to purchase the item. Weaker responses were frequently brief and usually did not link satisficing to the consumer then making a decision to purchase. Many candidates who did not seem to know the concept simply stated that the consumer would search until he/she is 'satisfied' with the product. This type of response was not creditworthy as it is restating the term from the question as the definition.

(b) There were many good responses to this question with some candidates showing a clear understanding of reductionism and being able to explain how satisficing is reductionist. Strong responses were able to explain why the concept is reductionist with a supporting example. Candidates who did well were often able to provide a further scenario where a customer does not use satisficing to purchase a product to illustrate their understanding. Many responses achieved 1 mark for this question. These responses were often brief and identified other consumer decision-making models such as utility or prospect theory. Candidates who achieved 0 marks in **part (a)**, were not able to achieve marks as they did not know what the term meant.

Question 7

(a) (i) There were many sound, full mark responses to this question with candidates having a good understanding of system 2 decision-making with an application to shopping for computers. Full mark responses identified system 2 as more conscious and deliberate thinking. Clear links were then made to buying an expensive item with many features to consider before purchase such as a computer. Weaker responses either were able to define system 2 decision-making without linking this to purchasing a computer or explained that computers are expensive items, without saying what system 2 thinking is. A very small number of responses either did not attempt the question or stated that computers are important to customers which was not creditworthy.

(ii) Similarly to **part (a)(i)**, there were many full mark responses with a clear definition of system 1 decision-making being quick/automatic. Most were also able to then explain why this type of decision-making is used by customers when shopping for groceries. Strong responses explained that grocery shopping is an everyday task and inexpensive. Many responses explained that customers frequently purchase the same groceries every week so they know which products they want without much thought. Weaker responses often could only outline that groceries were inexpensive, without saying what system 1 thinking is. A very small number of candidates either did not attempt this question or stated what groceries a customer might purchase which was not creditworthy.

(b) Most responses were able to give a brief outline of a problem such as being unable to observe what the customer is thinking or that customers might use both systems. These responses achieved 1 mark. Full mark responses included an example/scenario to support the reasoning. A few candidates did not attempt this question or stated that customers might not purchase the product which did not answer the question.

Question 8

(a) There were some Level 3 responses to this question. Many provided clear details of the features of menu design with both positive and negative impacts. Some responses summarised the features and others made reference to the Pavesic study on the psychology of menu design. Many candidates referred to the average time customers looked at a menu, eye-tracking, primacy-recency effect and eye magnets. Some candidates referred to common menu mistakes such as incongruence and overemphasising prices. Some outlined the Dayan and Bar-Hillel study on primacy, recency and menu item position. Weaker responses gave fewer positive and negative features. Some responses gave an anecdotal response with little to no reference to psychology of menu design. A common error shown in responses was to state that customers spend more time looking at the middle of the menu (and purchase more from here). From this, it appears these candidates were referencing the research into shelf position rather than menu item choice.

Responses were not as detailed for the second bullet point in this question on the effect of food name on menu item choice. Level 3 responses referred to the study by either Lockyer on the impact of menu item names on selection of menu item or Wansink et al. on the effect of descriptive food names on sensory perception in restaurants with clear details of the procedure and results of one of these studies. Weaker responses either gave fewer details of one of the studies or outlined that customers show a preference for the seasonal menu and the lowest to a French style menu. Some candidates gave anecdotal responses and outlined that customers prefer 'nice' or 'appetising' food names. These responses were not creditworthy.

(b) The best examples of strong evaluation went into detail about the named issue, generalisations, and made sure they referenced the research they described in **part (a)**. A very small minority used counter-points or gave an explanation for their analysis and often just stated that the studies were either 'similar' or 'in contrast'. Other common evaluation issues used were individual and situational explanations, applications to everyday life, and ecological validity. The best responses covered this question by structuring it by issues, along with supporting examples from the studies/features of menu design and the effect of food name on menu item choice from **part (a)** and analysis.

Popular responses for the named issue included how generalisation was low for the Lockyer study as it was done in New Zealand, or that the Dayan and Bar-Hillel study was only conducted in Israel. Candidates mentioned that the Dayan and Bar-Hillel study had good ecological validity as it was conducted in a café is a small town-centre coffee-shop. These responses discussed how this is a real life setting which reduces both bias and demand characteristics. Other strong responses included applications to real life, whereby candidates contextualised their response by providing examples from the study, and linked it to how it can be applied by managers in real life for restaurant profitability. Some responses remarked how the customer's influence of choosing preferences was situational due to the change in the menu/position of items but also individualistic due to a person's personal preference which showed good analysis.

Weaker responses were characterised by discussing the issue and applying it to one of the studies just by naming it. This evaluation was very superficial. For example, some responses just stated that a study had a poor sample due to being conducted in one country. These responses were often awarded Level 1. Some candidates attempted to use applications to real life but simply stated that the findings will 'improve sales' without elaborating how or providing evidence from the studies. A significant number of candidates simply repeated describing their response to **part (a)** which was not creditworthy.

Health Psychology

Question 9

(a) There were many strong responses to this question which were able to give specific ways to improve doctors' verbal communication. Popular responses included explaining medical terms in simple language/using less medical jargon. Some candidates further supported this with an example e.g. instead of saying myocardial infection say heart attack. Another common full mark responses suggested to use either the doctor-centred or the patient-centred approach. Those that achieved full marks explained what is meant by doctor/patient centred. One mark responses frequently either identified or very briefly outlined a way to improve doctors' verbal communication with patients. It was common for some candidates to answer **part (b)** of the question in **part (a)** with an explanation given for why the suggestion would be better for patients. No credit was given for this part of the response. Some candidates included non verbal examples such as using pictures or improving body language which were also not creditworthy.

(b) Responses that achieved 1 or 2 marks in **part (a)** were frequently able to achieve marks in this part of question. Good responses were able to outline why their suggestion in **part (a)** would be better for patients. Popular responses included explaining medical terms will help patients understand the doctor, and as a result help them feel more comfortable/relieved. Most candidates were able to identify why it would be better for patients, but few were able to outline the effect that it has on the patient which often achieved 1 mark. Some weaker responses focused their answer not on the patient, but on how it would benefit the doctor which was not creditworthy. Candidates who achieved 0 marks in **part (a)**, such as outlining how to improve non-verbal communication, usually achieved no marks to this question part.

Question 10

(a) The vast majority of candidates know what is meant by nature and nurture. Full mark responses outlined how nature is where behaviour is a result of genetics/is innate while nurture is where behaviour is due to learning/environment/surroundings. Some candidates confused the two terms and stated nature was due to environment and/or nurture was due to the genetics which was not creditworthy.

(b) Successful responses described how due to gate control theory suggesting that pain is caused by the environment or experiences within the individual (such as feeling stressed), then the patient can do something to reduce their pain. Most responses that received credit were able to identify that gate control theory gives the patient the ability to reduce pain. Some outlined that the distraction technique is helpful to close the gate and reduce pain. Quite a few responses were unable to identify the link between gate control and nurture, even when they had given the correct definition for nurture in **part (a).** Responses that just described how the gate control theory of pain was on the nurture side without saying how it is a strength were not creditworthy.

Question 11

(a) There were many strong responses to this question. Candidates effectively included examples of using acupuncture and TENS, gaining full credit because they were also able to show how it works and then how it reduces Becky's pain. For example, linking the release of endorphins which reduce pain sensations. For acupuncture, Qi/life force is out of balance and acupuncture restores the flow of Qi so that the pain from the broken leg is reduced when small needles are inserted into the area where pain is experienced. Weaker responses could only include the procedure of acupuncture and TENS and were not linked to how these alternative treatments would reduce Becky's pain. A significant number of candidates suggested biological and/or psychological treatments such as imagery, paracetamol, relaxation, etc. which were not creditworthy. There were also some treatments suggested that would not have any effect on Becky's pain such as ECT and mirror therapy which is used for phantom limb pain. Neither of these treatments could receive credit.

(b) Some responses to this question gave a clear weakness of either acupuncture or TENS therapy and this was usually to state that it was time-consuming or expensive with either an example to back up this weakness or a brief explanation as to why. A small portion of candidates mentioned that TENS only provides relief for a short amount of time, which results in the patient having to go back to the clinic.

Many of the candidates did not seem to know a weakness for these alternative treatments as there were a number of very brief responses that did not achieve any credit. Some stated that either acupuncture or TENS are not effective as they are not well researched which is incorrect as a number of studies have been done on effectiveness of both. Some of the research shows it is effective and some does not. Another type of non-creditworthy point that was common was to exaggerate the ethical issues with acupuncture, for example, that it is extremely painful. It may cause some discomfort but this is not very painful and the practitioner would stop the treatment if it became too much the patient.

Question 12

(a) The responses to this question covered the full range of the mark scheme. Stronger responses gave clear and often detailed and accurate description of unrealistic optimism and positive psychology. Strong responses outlined study 1 by Weinstein on unrealistic optimism in college candidates. A few also attempted to outline study 2 although this often was not in detail. It should be noted that an outline of one of the studies is sufficient for a Level 3 mark. Some responses outlined in detail the three features of positive psychology – pleasant, good and meaningful life. Other responses either outlined the study by Seligman or the study by Shoshani and Steinmetz on positive psychology at school. Some responses attempted to outline all three features of positive psychology and both the studies by Seligman and Shoshani and Steinmetz. These responses were often over-written which left candidates little time to respond to **part (b)** or all three were discussed very briefly which led these responses to achieve in the lower mark bands (frequently Level 2) due to lacking detail or showing good understanding. Similarly to unrealistic optimism, one of the studies or a detailed outline of the three features of positive psychology would achieve Level 3.

Responses achieving in Level 1 often lacked detail and were sometimes inaccurate. Level 1 response frequently just defined unrealistic optimism and identified the 3 concepts in positive psychology without any definitions or examples.

Common responses that were not creditworthy described that unrealistic optimism is having irrational thoughts or that people with unrealistic optimism do not suffer from mental health problems. Positive psychology is changing unrealistic thoughts to more positive thoughts and both unrealistic optimism and positive psychology are used by doctors to help people improve their physical health.

(b) The marks awarded to responses to this question were varied with many achieving in Level 1 and Level 2 due to a lack of specific examples and no or very limited analysis. Stronger responses that achieved Level 3 and above were structured issue-by-issue and covered the named issue of nomothetic versus idiographic approach with examples from the concepts and/or studies described in **part (a)**. For the named issue, candidates mentioned how the use of questionnaires in the studies came from a nomothetic approach, and that at the same time, the unrealistic optimism study is also idiographic as participants had to rate their personal experiences. In addition, strong responses discussed how positive psychology is nomothetic as everyone will want to achieve a pleasant, good and meaningful life. However, it can also be seen as idiographic as what gives one person happiness, good relationships and meaning in life will be unique to each person. Other common evaluation issues included determinism versus free will, generalisability and applications to real life on how the use of positive psychology can be used to improve well-being of other people.

Weaker responses often discussed the issue and applied it one of the studies or concepts from **part (a),** frequently by just naming it. This evaluation was very superficial. Other weaker responses talked about how the studies were standardised and this improved reliability but no context was given which limited the marks awarded (frequently Level 1). Weak responses often gave incorrect definitions of both idiographic and nomothetic. For example, stating that idiographic is whether something can be applied and used in the real world which is not creditworthy.

Organisational Psychology

Question 13

(a) Responses were varied for this question with some outlining a clear suggestion that could monitor employees working from home. Common creditworthy responses included using technology to monitor worker productivity such as a webcam or electronic performance monitoring (EPM). Strong responses were able to make the link to how it would monitor the workers e.g. for EPM – allow Feba to track their mouse clicks which would be able to monitor their productivity. Some candidates missed the fact that Faba is an IT manager and therefore technical equipment would be the only realistic way to measure productivity. Responses that were not creditworthy as they did not measure productivity included asking staff to write lists of what they were going to do or have a meeting at the end of the week to discuss the work.

(b) Some responses to this question achieved full marks by explaining how either social facilitation theory or the Hawthorne effect due to the monitoring of work would increase their productivity. A few responses linked to the study by Claypoole and Szalma which used Electronic Performance Monitoring and concentration levels and this produced good responses so long as the response remained focused on the increase in productivity due to the monitoring. Weaker responses were often due to giving an incorrect response in **part (a)**. For example, if the candidate outlined giving staff lists of what they are going to do or having a meeting at the end of the week, they found it very difficult to explain why this would increase productivity and instead stated they would 'work hard' to look good to their manager which was not creditworthy.

Question 14

(a) There were some strong, full mark responses to this question. These candidates were able to give a clear definition of holism and then relate it to Thomas-Kilmann's five conflict handling modes (the fact that it takes into account 5 modes show that it covers many possible solutions). Some responses achieved 1 mark by giving an outline of holism but could not give an example from Thomas-Kilmann's five conflict-handling modes that illustrates holism. Candidates who did not

achieve any marks for this question were not able to define holism and sometimes explained it was connected to ecological validity or practical applications.

(b) Stronger responses to this question were able to explain one weakness and put this into the context of the Thomas-Kilmann's five conflict-handling modes. Common weaknesses included that it can be time consuming or requires a lot of effort/training to identify which of the five ways are being used to resolve the conflict. It was common for responses to attempt to explain why Thomas-Kilmann can be considered to be holistic which was not creditworthy.

Question 15

(a) (i) Full mark responses identified that the error was an error/sin of omission and linked this clearly to Joseph forgetting to wear eye protection which he should have remembered. Weaker responses often just identified that it was an error of omission. Some responses restated that it was a human error which was not creditworthy as this was in the question. In addition, some incorrectly identified that it was an error of commission.

(ii) Many responses achieved 1 if not 2 marks for this question by suggesting a specific idea of how the management at this car factory could avoid this error of omission about eye wear. Common suggestions included putting up posters, having a checklist for safety, using token economy to reward workers that wear the eye protection or carrying out training to ensure workers know the proper safety equipment required. Other popular solutions were assigning someone to spot-check each person's safety gear before work. Weaker responses were usually due to the candidate not linking this to wearing the eye protection. Suggestions that were not creditworthy included stating that the factory needs to implement safety measures or that workers must put on eye protection which is something that the factory already does.

(b) Many responses were able to offer a brief weakness for the suggestion given in **part (a)(ii)**. Common responses included that workers may not pay attention to the signages/time consuming/costly to employ someone to spot-check, or that it would be costly to implement token economies and rewards. Many responses did not refer to protective eyewear which limited their mark on this question to 1. Responses that did not receive any credit often gave an unclear weakness, such as it would be difficult to do or that it would not work.

Question 16

(a) There were a range of responses to this question covering the full range of the mark bands. Many responses gave clear and detailed descriptions of Maslow's hierarchy of needs and McClelland's theory of achievement motivation. They were able to give full descriptions of the theories and link to the workplace as they went through. A small number of responses were able to bring in the Saeednia study which was often detailed.

Level 2 responses frequently lacked detail and were sometimes inaccurate in terms of some of the definitions of the stages in Maslow's hierarchy and/or some of the three needs from McClelland's theory. Some of these Level 2 responses did not link one or both of theories to work in their response. Level 1 responses sometimes just outlined one of the two theories in a bit of detail or gave a brief outline of both. It was common for these responses to just list the three needs and then the five or eight parts of Maslow's hierarchy.

Responses that did not achieve marks for this question, sometimes just identified one of Maslow's needs and then stated McClelland's theory is where motivated workers want to achieve.

(b) The marks for this question tended to be between Level 1 and Level 3. A small number of responses achieved Level 4 and above by giving detailed examples from **part (a)** and presenting analysis throughout the response. For the named issue, candidates who covered the issue well mentioned about how Maslow's theory is deterministic as the hierarchy is fixed and the same for everyone. In addition, workers can only move up a level when the needs of the previous level have been met. Some responses were able to provide analysis for this theory by discussing the free-will side of the theory. Workers will have some influence on the requirements for the level of the hierarchy. For example, employees will have chosen what level of social support they need at work. Strong responses were also able to identify that the three needs from McClelland's theory are deterministic as all workers need to have these three needs met. However, there is also an element of free will as workers will choose which one or two needs they value above the others.

Other common issues included cultural differences as the theories were westernised. Some responses discussed that people in collectivistic cultures would have different needs or that the levels may be ordered differently for different cultures. In addition, application to everyday life, reductionism versus holism and individual versus situational explanations were popular issues for this section.

Weaker responses often identified and sometimes defined some of the issues. The response then discussed the issue and applied it one or both of the theories from **part (a),** frequently by just naming it. Sometimes Level 1 responses evaluated both theories together. It was typical for these candidates to state which side of the debate they thought both theories fell on e.g. 'they are both holistic' or 'they are both deterministic'. This evaluation was very superficial. These responses might state that the theory 'only' considers 3 needs or 5 levels.

PSYCHOLOGY

Paper 9990/32

Specialist Options: Approaches, Issues and Debates

Key messages

Questions 1, 3(a), 5, 7(a), 7(b)(i), 9, 11(a)(i), 11(a)(ii), 13, 15(a)

These questions in this exam asked candidates to apply an area of the syllabus (theory, technique/treatment, self-report etc.) to explain how it is relevant to a particular scenario or context. It is important that candidates are aware of the titles of the bullet points in the syllabus. It would be helpful for candidates to write revision notes with the title of the topic area and bullet point at the top so that they can identify which part of the syllabus these types of questions are referring to. Candidates should also refer directly to the scenario/context in the question in their response.

Questions 3(b), 7(b)(ii), 11(b) and 15(b)

These questions in this exam asked candidates to evaluate the suggestion such as the technique or treatment that was outlined in the candidate's response to **part (a)** of the question. In this exam, these questions asked the candidate to evaluate the technique outlined in **part (a)** such as with a weakness, explain why the self-report in **part (a)** is valid or identify a problem with the technique outlined in **part (a)**. When revising, it would be helpful to candidates to learn strengths and weaknesses of the theories, techniques, self-reports, treatments etc. they have learned and put these into their revision notes. They should also practice explaining the evaluation point in the context of the question.

Questions 2, 6, 10 and 14

Part (a) – These questions could ask the candidate to outline a theory, study, technique/treatment or self-report used by psychologists that is named in the syllabus or outline one of the issues and debates, possibly with an example from the syllabus content. The revision technique outlined previously in this report will aid candidates in learning the syllabus material.

Part (b) – This part of the question may ask candidates to explain a strength or a weakness of the issue/debate or the syllabus content outlined in **part (a)**. The question could also ask candidates to explain how a bullet point in the syllabus links to or supports one of the issues or debates. It would also be useful for candidates to write revision notes where they define the issues/debates and prepare a strength and a weakness of each issue and debate to prepare for **part (b)** of this type of question. Candidates should also note how the topics covered in the syllabus fit with each of the issues/debates. These questions were worth 2 marks for each part of the response and therefore a short response is appropriate.

Questions 4(a), 8(a), 12(a) and 16(a)

These questions in this exam came from one or two of the bullet points in the syllabus. Candidates can be asked to outline a key study from the syllabus or two studies, theories, characteristics/explanations/treatments 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 could be useful for candidates to create revision notes with the title of each topic area and the description in the bullet point as the header. Alternatively, candidates could create a mind map and put this information in the centre.

Questions 4(b), 8(b), 12(b) and 16(b)

These questions will always ask the candidate to evaluate the studies, theories, characteristics/explanations/treatments of disorders or techniques described in **part (a)** of the question. The response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. Most responses that evaluated using two issues in this exam limited their marks as the response tended to be 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 studies, theories, characteristics/explanations/treatments of disorders or techniques described in **part (a)** of the answer.

Candidates must also provide some form of analysis to access Level 3 and above. 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. The response needs to explain the comparison/strength/weakness or counter-argument with examples from **part (a)** of the question. It was common for responses to state that two theories, for example, were 'similar' or 'in contrast' for an issue without any explanation as to why they could be compared in this way. This is limited analysis. 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 4 and 5 descriptors, it would be best to structure the response by issue rather than by study and/or theory. It would also be recommended that the response starts with the named issue to make sure the answer covers this requirement of the question.

A small minority of candidates did not evaluate using the named issue. Quite a few of the answers were structured by study/theory/treatment rather than by the issue which often led the response to be quite superficial and repetitive. Candidates should be aware this question is worth 10 marks and needs to include an appropriate amount of information.

General comments

The marks achieved by candidates for this session of the 9990 specification achieved across the full range of the mark band. Stronger responses showed good knowledge, understanding, application and evaluation. Weaker responses showed limited knowledge and understanding with brief, superficial and sometimes anecdotal answers. These responses often had limited evaluation and application skills.

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

The questions on clinical were the more popular choice of option, followed by health.

Comments on specific questions

Clinical Psychology

Question 1

The responses to this question covered the full range of marks. Common responses included reference to tensing muscles and preventing fainting due to raising blood pressure. Stronger responses often referred to the timing of tensing and relaxing muscles (e.g. 20 seconds tensing and then 20 seconds relaxing) and suggested that Craig practice this multiple times per day and use it if he feels anxious or faint when watching television. Weaker responses were sometimes confused with the Delapp study and outlined other treatments. Some said 'use applied tension' without explaining what this is. It was common for responses to not explain how applied tension could help Craig watch television which limited the mark awarded to a maximum of 3. A significant number of responses outlined the incorrect treatment such as systematic desensitisation, imagery and progressive muscle relaxation which were not creditworthy.

Question 2

(a) Most responses achieved 1 mark for outlining what is meant by nurture. They were of an appropriate length and gave short, clear answers describing that nurture refers to behaviour that is learned or environment influences behaviour. Many responses were also able to give a comprehensive outline of the behavioural approach for a fear related disorder including the use of the key terms which were explained appropriately, such as 'pairing' or associating a neutral stimulus with an 'unpleasant/fearful' one. Weaker responses wrote about operant conditioning as an example or identified classical conditioning without any explanation of what it means so could not achieve the second mark. A few responses attempted to outline 'nurture' but just stated this is how a person is 'nurtured'. As this uses the same wording from the question to define the term, it is not creditworthy.

(b) There were a number of strong responses to this question that achieved full marks. Candidates could often identify a weakness but did not then apply this to fear-related disorders. Reductionism was a commonly used response, with many stating that it ignored biological or genetic influences. Another correct and frequent response was how some phobias are easier to acquire than others as there was a greater genetic predisposition towards them e.g. it is easier to learn to be afraid of snakes than balloons. Responses that were not creditworthy identified a weakness of the Little Albert study, such as ethics and so seemed to misunderstand the question. Weaker answers identified an issue/debate, such as reductionism, but did not explain how it is reductionist by missing out other explanations.

Question 3

(a) Responses to this question covered the full range of the mark scheme. Those that achieved 3 – 4 marks often included many features of the BDI. Most stated that it could be used with Joan to aid diagnosis and that it had 21 items. Many responses could also give the different levels of depression from the test scores with some accuracy. Some responses identified 0 – 3 scale and related this to the symptoms experienced by Joan in the question, such as her sleep problems.

There were some inaccuracies about the BDI itself. A few described Beck's theory rather than the measure. Weaker responses gave incorrect details about the scale, number of items and the range of results that indicate different levels of depression. A significant number of responses did not link in one of Joan's symptoms from the context. Candidates should be advised to use the context provided in the question for application questions. Some responses stated that the BDI was a treatment and why it would be used which was not creditworthy.

(b) There were some full mark responses to this question that could give a clear reason why the BDI is valid. The words 'concurrent validity' and 'correlated with' other measures were regularly and appropriately used. The Hamilton Depression Scale was the frequent example of the other measure. A very small number of strong responses referred to the BDI being closely linked to the symptoms and diagnostic criteria of depression in the ICD-11. Many responses confused validity with reliability so explained that it produces a quantitative measure and can be re-tested or comparisons can be made. This type of response was not creditworthy.

Question 4

(a) Responses varied for this question and covered the full range of the marks available. Level 3 responses often gave clear details of electro-convulsive therapy (ECT) and cognitive-behavioural therapy (CBT) for the treatment and management of schizophrenia. Stronger responses referred to seizures, voltage used, anaesthetic, and unilateral application of electrodes for ECT. A few responses also described how ECT can help reduce positive symptoms and/or is frequently used for severe cases of schizophrenia. Strong responses also included details specifically linked to schizophrenia, including change of irrational thinking, sometimes with a specific example that someone with schizophrenia might describe to their therapist and details of the more rational explanation the therapist could provide. Many successfully referenced the Sensky study.

Weaker responses tended to be too general or not sufficiently focused to be creditworthy. For example, they described the process without any specific detail beyond 'electrodes to the 'head' or 'a talking therapy' respectively or referred to the history or side-effects of ECT or the steps of CBT without stating what/how the treatment contributed to the management of the schizophrenic symptoms specifically. There were some misconceptions that candidates need to be aware of in

terms of ECT. Some responses referred to side effects which was not creditworthy in **part (a)** of this question.

(b) The marks for this part of the question covered the full range of marks available with the most frequent levels awarded being Level 2 and 3. Those that achieved Level 3 and above structured their response issue by issue, and often started with the named issue of idiographic/nomothetic. For example, a Level 5 response could state ECT is nomothetic because of the universal/standardised procedure as well as being slightly idiographic due to the more extreme cases receiving ECT than less severe cases. Similarly, CBT is nomothetic in terms of the steps and process but idiographic as the symptoms given and the challenges to any irrational thoughts will be unique to each schizophrenic patient. Responses achieving Level 4 would have given one of the two comparisons shown above and Level 3 would be less likely to argue that both treatments can be seen as nomothetic but would state there is a difference with ECT being more nomothetic and CBT being more idiographic. Other common issues included nature versus nurture, determinism versus freewill, reductionism versus holism and ethics. There was some misunderstandings in terms of the ethics of ECT with many candidates thinking that broken bones and death are frequent outcomes of ECT which is incorrect. It is also difficult to calculate if the reason for death is due to the ECT or the administration of anaesthetic. Research has shown death occurs 2.1 times per 100,000 ECT treatments. Nowadays as lower doses of electricity are used, broken bones are very rare. Memory loss is a creditworthy side effect. It should be noted that for most patients memory loss is resolved within two months of treatment (i.e. the memories return). Only 7 per cent of patients have memory loss after 12 months.

Weaker responses demonstrated some confusion between the nomothetic and idiographic approaches though others could define them but did not effectively demonstrate them using the details of the respective treatments. When a large variety of evaluative issues were included, responses tended to repeat the details from other points they had used previously. Other weak responses proceeded to extend on the description they had included in **part (a)** which was not creditworthy. Others listed a number of 'explained' evaluative points and gave no corresponding evidence from the treatments described in **part (a)**.

Consumer Psychology

Question 5

There were a number of strong responses to this question and many achieved full marks. These responses outlined design features from each menu and this was contextualised with the mushroom pizza and how this would increase its sales. Use of an 'eye magnet' and putting the mushroom pizza on the 'top or bottom' were the most used design features. Weaker responses did not specifically connect the design feature to mushroom pizza and stated the benefits of a particular design feature in simple terms. It was rarer to see candidates explain how the primacy/recency effect improved memory for the mushroom pizza. A significant number of responses stated that the mushroom pizza should be put in the centre of the menu. These responses were confusing appropriate menu design features with shelf position and were not creditworthy.

Question 6

(a) Many responses were able to give a clear outline of both individual and situational explanations. The terms 'innate' and 'personality' were used for individual and 'environment' was used for situational explanations to gain full marks. Candidates were more likely to get credit for situational explanation definition rather than individual. Some responses referred to the words in the question without any further explanation, for example, individual explanation refers to the individual which was not creditworthy.

(b) Some responses were able to gain full marks. These responses were able to give a definition of overload that was clear. In addition, responses alluded to the presence of another in the environment to be stressed/overwhelmed which clearly showed the situational aspect. Weaker responses did not correctly outline what overload meant or explain why it is situational. Very few explained that overload is related to too much information.

Question 7

(a) There were some sound, full mark responses to this question. Common responses included multiple unit promotion with a simple example of a possible promotion related to tins of tomatoes,

for example, 3 for $2. Quantity limits and other reference to anchoring were much less common. Weaker responses were often anecdotal or incorrect, with no use of key terms or made a general suggestion without any example to support it. Many started to explain the thinking behind it without realising that was the focus in **part (b)(i).** There were some incorrect suggestions such as instead of a tin of tomatoes for $1 you could have two tins for $3.

(b) (i) Full marks answers included what customers thought and why they would buy more tomatoes. For example, 'They would think they were getting a discounted price on each tin of tomatoes if they bought more so therefore bought more in order to access the discount'. Other common responses included when anchoring was suggested in **part (a),** that customers would think about making the dish suggested (e.g. buy 4 tins to make spaghetti bolognaise) and would therefore want 4 tins to make it. Weaker responses were brief and often stated that the customer would think the tomatoes were cheaper without linking this to why this would make the customer purchase multiple tins rather than a single tin of tomatoes. If the candidate gained 0 marks for **part (a)**, they frequently did not get any marks for this question.

(ii) Most responses were able to give a brief outline of a problem which were often practical rather than psychological. The most popular answer was around people not buying in bulk because they lived alone and so did not need multiple cans of tomatoes and so no offer would not work with them. In addition, there were a significant number who pointed to strategies that implied a discount (when there was not one) might actually put customers off returning to the store as they had been duped. Weaker responses were often brief such as stating that some customers do not want more than one tin without any explanation as to why this might be the case. Overreaching reasons doubting the quality of the tomatoes were provided as a potential problem, which was not relevant to the point purchase decision strategy suggested and therefore was not creditworthy.

Question 8

(a) There were many strong Level 3 responses to this question. These responses included the three main requirements of the answer: the independent variable, the sample and the results. Many responses also included details about location, type of restaurant, design, number of nights and controls used. Weaker answers often gave vague details e.g. large sample, a field study without saying what/where and the overall finding regarding classical music. The results were often very general for Level 1 and Level 2 responses and could have been more specific regarding spending in the classical music condition as compared to the other two conditions.

(b) There were many Level 3 and above responses to this question. Most responses evaluated using the named issue of ecological validity with many giving some very clear examples from the North et al. study. Analysis was given in some responses with a discussion of the advantages of the study being done in a natural setting as well as the disadvantage of extraneous variables. Other common evaluation issues included generalisability, determinism, reductionism, application to everyday life and quantitative data.

Weaker responses lacked depth in their discussion of ecological validity with a few confusing ecological validity with other types of validity. Some responses made points without referencing the necessary detail from the study, i.e., the ecological validity was high because it was a field study. There was a good possibility of developing the generalisability of the sample by providing context and analysis between different cultures which was not well developed by most candidates. Often responses had no links to the study and were just generic points. Candidates should think about the issues that work best. Many wanted to include nature nurture, when situational and individual was more appropriate. Weaker responses tended to provide a long list of points with a vague connection to the study.

Health Psychology

Question 9

There were many strong responses to this question. These responses handled each aspect of the Health Belief Model in turn with an appropriate suggestion, e.g., perceived susceptibility is an important aspect to consider as if people do not think the issue will affect them, they are less likely to change their behaviour. Therefore, Dr Munsi should explain in her leaflet that age does increase one's susceptibility to heart disease and the complications that arise from it so that they realise the importance of taking their medication every day and/or the risks of not taking their medication every day. In addition, many gave some detail as to what

the consequences might be, such as hospitalisation. Stronger responses made specific reference to heart disease and the demographic of the group (elderly patients). Weaker responses gave details of the Health Belief Model without providing specific ideas which Dr. Munsi could include in the leaflet which were informed by the model. In addition, some responses suggested fear arousal and did not link it to the Health Belief Model or made general suggestions with no mention of the Health Belief Model. Some candidates missed that it was a leaflet that needed to be produced and suggested meeting with the patients or giving them pill counters without referring to a leaflet at all. These responses were not creditworthy.

Question 10

(a) Many responses were able to outline what is meant by 'application to everyday life' by describing the usefulness or how it could help people in the real world. Many responses also outlined a measure of non-adherence with TrackCap and pill counting being the most common choices. Responses that did not gain any credit included outlining the term as meaning 'it can be applied to daily life', which does not answer the question as it is using the same wording as the question. Some measures were mentioned without specific link to adherence. Some responses did not include a measure of non-adherence.

(b) There were many strong responses to this question. These responses went beyond simply mentioning limitations of pill counts. They discussed how factors like patient dishonesty due to social desirability or confusion about medication instructions can compromise the validity of self-reported measures. This demonstrates a clear understanding of how inaccurate reporting can impact our ability to assess actual medication adherence. Another common response included that patients would take the medication out and not take them. Weaker responses often discuss problems in the validity of medication adherence measures by citing the example of pill counts/TrackCap and stating patients might lie about taking medication or remove pills without consuming them. This response, while acknowledging a potential problem, lacks depth and demonstrates a surface-level understanding of the issue.

Question 11

(a) (i) There were some strong responses to this question. Full mark responses identified positive emotions and were able to identify appropriate pleasurable activities for an 8-year-old such as playing, baking, going to the park, etc. Many candidates did not know what a 'pleasant life' referred to and instead gave a suggestion for either a good life or meaningful life or did not give specific advice. These responses were not creditworthy.

(ii) There were some strong responses to this question. These responses used key terms such as service to others, larger purpose and institutions. Specific advice like charity and volunteering were provided. Similarly to **part (a)(i)** although less frequent, candidates did not know what a 'meaningful life' referred to and instead gave a suggestion for a good life or did not give specific advice. These type of responses were not creditworthy.

(b) Some responses to this question were able to gain full marks by clearly explaining one reason why using positive psychology might not lead to an improvement in mental health for Zainab's daughters. The most common successful responses related to either spending time on these activities being stressful and therefore not helping or that the daughters might have mental health issues which might prevent them from benefitting from the positive psychology. Weaker responses were often very brief and just outlined a problem without clearly linking it to positive psychology or the age of Zainab's daughters. Some responses stated that her daughters would not want to do the suggestions or that they need to read a book on positive psychology. Neither of these responses were creditworthy.

Question 12

(a) The responses to this question covered the full range of the mark scheme. Stronger responses gave a clear and often detailed and accurate description of a study investigating reasons for delay in seeking treatment and a study on Munchausen syndrome. These responses often gave clear details of the study including the sample and self-report used. They also gave clear definitions of the three types of delay. Many described the impact of levels of pain, social stigma, fear of diagnosis, or cultural beliefs on delaying treatment seeking. Additionally, these responses considered the consequences of delays on health outcomes and healthcare costs. The case study for the Munchausen syndrome was well detailed in most Level 3 responses. The case study by

Aleem and Ajarim was the most popular choice with very few responses choosing a different case study. Weaker responses often gave fewer details about delay in seeking treatment and would outline the three types of delay but sufficient details of a study was largely not presented. The case study on Munchausen syndrome for weaker responses lacked detail or focused on Munchausen syndrome symptoms without describing the case study.

(b) Responses to this question covered the full range of marks. Stronger responses effectively evaluated the case study method. They recognised the ability of case studies to generate rich qualitative data, offering valuable insights into conditions like Munchausen Syndrome. Additionally, these responses acknowledged the limitations of generalisability inherent to studying a single case and usually gave examples from the Aleem and Ajarim study. Other common issues/debates used well included application to everyday life, type of data, idiographic versus nomothetic.

Weaker responses attempted definitions of the evaluation issues with some success. The response then discussed the issue and applied it one of the studies or concepts from **part (a),** frequently by just naming it. This evaluation was superficial. Some responses displayed a limited understanding of case study evaluation, such as stating that case studies are not generalisable because they involve one person, without elaborating on the implications of this limitation for the broader understanding of a condition like Munchausen Syndrome. Additionally, these responses might have solely focused on limitations without mentioning any potential strengths of the case study method. Points in these weaker responses were either fairly simple or incorrectly applied. For example, there was a tendency that if one study was argued to favour on one side of a debate then the other research should be argued the other way, such as Safer's research was holistic considering several factors that contributed to the delay in seeking advice so that the Munchausen research was argued as reductionist because it only considered one individual.

Organisational Psychology

Question 13

Responses varied for this question with some able to give clear reasons why Tia was bullying Arjun using appropriate knowledge of the causes of bullying at work. Responses that gave two suggestions tended to be the ones that were most likely to achieve full marks. Stronger responses handled each 'origin of bullying behaviour' in turn with an appropriate suggestion, e.g., an individual characteristic, either personality/status with the bullying behaviour and then identify a second, different reason, using the correct terminology, e.g., scapegoating and relate it appropriately to the stimulus material. Stronger answers also recognised the factor that this bullying was taking place in the workplace. Some responses displayed a limited understanding of the reasons behind workplace bullying. They might have mentioned Tia's promotion and stress but failed to explain how these factors could lead to bullying behaviour. Additionally, these responses might have offered a generic explanation for bullying that did not address the specific context of the question.

Question 14

(a) There were some strong, full mark responses to this question. These candidates identified the 2 factors and gave appropriate examples. There were many examples of successfully described hygiene and motivational factors. These included examples as well as the fact that they were different in terms of whether they lead to satisfaction or preventing dissatisfaction. Responses that did not receive credit for this question incorrectly identified the 2 factors or got them the wrong way round, with inappropriate examples of the factors when present.

(b) Stronger responses effectively explained the applicability of Herzberg's theory across cultures. They recognised the adaptability of the theory due to its focus on broad categories without strict definitions of motivators and hygienes. Therefore, different cultures could have different definitions of what they see as motivators and hygienes. Additionally, these responses might have suggested strategies for applying the theory in different contexts, such as identifying culturally relevant motivators and hygiene factors. Weaker answers often just repeated the question, that it could be applied across cultures, without explaining why or were generic/anecdotal so it was hard to credit a reason. Some responses argued that it could not be applied to all cultures which was not creditworthy.

Question 15

(a) There were a number of strong responses to this question. The vast majority of candidates were able to suggest at least one correct job design technique, if not two. Stronger responses effectively explained job design principles and their application to improve job satisfaction in a toy factory. They identified specific features like job rotation, explained their purpose, and provided relevant examples of how rotation could be implemented in a toy factory setting. These responses demonstrated a clear understanding of how job design can be tailored to a specific work environment. Other strong responses included suggesting job enlargement and sometimes job enrichment. Increasing autonomy was also well applied. Weaker responses were often due to no clear link to either the toy factory or the impact on employee satisfaction even though two job designs were identified. Use of intrinsic and extrinsic motivation was often wrongly suggested as these are not job design techniques and therefore not creditworthy.

(b) Many responses were able to offer a brief weakness for one of the job design techniques suggested in **part (a)**. Stronger full mark responses effectively discussed a weakness of a job design techniques. They identified a weakness for job rotation and explained how the need to constantly learn new tasks could be stressful for workers, potentially reducing job satisfaction. Other examples of full mark responses included that job enlargement might cause resentment at having to do more and job enrichment might not work if the upskilling did not work and the responsibility was more than the individual could cope with, resulting in demotivated workers. Some responses displayed a limited understanding of the weaknesses associated with job design techniques. They might have mentioned job rotation as a problem but failed to explain why learning new skills could be stressful or decrease job satisfaction. Some responses were focused on an incorrect suggestion from the first part, for example increasing pay, and thereby 'costing the company more money'. Some candidates did not attempt the question (if the first part had not been attempted).

Question 16

(a) There was a range of responses to this question covering the full range of the mark bands. Many responses gave clear and detailed descriptions of token economy used to reduce accidents at work and monitoring accidents and risk events. Most responses described the Fox et al. and Swat studies in some depth. Details for the Fox et al. study included where the research took place, that it was a longitudinal study, the conditions the tokens were issued (and not), and some detail regarding the data about the reduction in days taken off for injury. Likewise, for Swat's study many candidates demonstrated good mastery of the material (where, what and how the study was conducted) and understanding of the results, what accidents were more frequent/severe than others, the different types of accidents that took place in the different factories/workplaces and the contributing factors that led up to the accidents occurring along with the suggestions which should be put into place.

Weaker responses explained a token economy without any reference to the Fox et al. study. Although this was not a requirement of this question, responses that did not describe the study were often very brief in the details of how token economies would reduce accidents at work. Similarly, the Swat research was covered with varying (though minimal) degrees of detail about where, what, or how the research was conducted and where evident, possibly one result mentioned. For example, most accidents were caused by bad housekeeping or the most frequent accidents were in the 'slips and falls' category and mostly in the meat packing factory. Some candidates attempted part of the question with the most common one chosen being token economies. These responses were limited to Levels 1 and 2.

(b) The marks for this question tended to be between Level 1 and Level 3. Candidates could correctly identify the objective and subjective data in each study and provided examples, e.g., in Fox's study, objective data included the records pertaining to the number of days a worker or his/her team had taken off as a result of an injury and subjective data included the anecdotal data pertaining to the wife driving 50 kms to take advantage of the tokens they had earned before the initiative was going to be stopped. Similarly in Swat's study, the percentage of accidents that occurred in the different categories from the 83 accidents recorded and investigated as well as the subjective data generated from the interviews with the supervisors and floor managers on their perception of the causes of the accidents. Other evaluative issues that were successfully applied included generalisations, applications to everyday life, longitudinal method (particularly with reference to Fox et al.) and validity (which covered ecological validity, longitudinal method, and the data collection methods).

Weaker responses sometimes confused the objective/subjective data incorrectly as qualitative/quantitative data. Often, points were listed without being contextualised and when the idiographic and nomothetic approaches were included (or reductionism versus holism), there was usually repetition from the objective/subjective data points made earlier. Individual and situational was a relatively common choice but often contained misconceptions. For example, some candidates claimed that token economies were more individual because the individual could decide whether to be influenced by them or not, in spite of the research showing that most of the employees in the study were influenced by them. Similar confusion arose over determinism and free will where a case was made for free will for the same reason.

PSYCHOLOGY

Paper 9990/33

Specialist Options: Approaches, Issues and Debates

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

PSYCHOLOGY

Paper 9990/41

Specialist Options: Application and Research Methods

Key messages

- What has been learned from the AS component of the syllabus should be transferred to the A2 component. For example, at AS candidates learn about methodology, such as experiments, which also apply to A2.
- Questions should be read carefully ensuring that the focus is on what the question asks.
- For ***Section A*** answers, candidates should relate their answer to the study in question or include an example. Questions frequently end with 'in this study' and so the answer should be related to that specific topic area/study.
- All terminology should be explained. Writing 'it is valid and reliable', for example, is insufficient without explanation, application or example.
- The syllabus includes for 'example studies' such as 'e.g., Oldham and Brass (1979)'. Example studies can be substituted for alternatives, but these alternatives must cover the same or very similar content to the example study. If the Oldham and Brass study is substituted, the alternative study must be about a move to open plan offices and the data that was gathered from that move. The alternative cannot be about something different.

General comments

Some candidates answered questions from one option only. Other candidates, who correctly answered two options, sometimes performed considerably better in one option than the other.

Many candidates answered two questions from ***Section B*** instead of one (only one of these ***Section B*** responses can receive credit). Candidates are advised to read the instructions on the front cover of the question paper and to read the heading instructions for each question section.

Candidates should double check that the terminology they use in their answers is correct. Often terms such as reliability and validity were used interchangeably, as were qualitative and quantitative, and independent and dependent variables. There was also confusion with the terms format and technique in relation to questionnaires and interviews.

Section A

Question **part (c)** requires a general evaluative point that could relate to any study (such as a strength or weakness of a method) but it also requires for the general point to be related to the specific sub-topic/study in the question. Answers often included strengths and weaknesses but these were often not related to the question, and so restricted marks.

Candidates should not use psychological terms without explanation. Frequently, answers were limited to 'it is reductionist' or 'it is useful in everyday life' without further explanation. Stating 'it is reductionist' does not automatically identify it as a strength or weakness.

Candidates should not use the terms reliability and validity to answer every **part (c)** question for three reasons: **(i)** they do not apply to most questions and so cannot be awarded marks, **(ii)** candidates using the terms often do not know how they apply to the specific question and **(iii)** candidates often confuse the terms.

Section B

Candidates should only answer one question from this Section.

Many candidates appeared to assume that they must conduct an experiment whatever the question. An interview, questionnaire or observation are methods independent of an experiment and candidates are advised not to try to make other methods 'fit' into an experimental format.

Some candidates evaluate their plan in **part (a)** by listing strengths and weaknesses. This should not be done because: the question does not ask for evaluation; there are no AO3 marks allocated to evaluation; evaluation is done in **Questions (c)(i)**, **(c)(ii)** and **(c)(iii)**.

Some candidates included a paragraph of results. This achieves no marks because the question asks for a plan only. Further, the proposed plan has not been carried out, so no actual results are gathered.

Candidates need to know the distinction between questionnaire format and technique, and interview format and technique, as stated on the syllabus: Questionnaire technique: paper and pencil (i.e., done by a person with the researcher present), online or postal. Questionnaire format: open and/or closed questions. Interview technique: telephone or face-to-face. Interview format: structured, semi-structured, unstructured.

When using psychometric tests, candidates should not use acronyms unless the full title of it is provided first. For example, 'Beck Depression Inventory (BDI)' is fine, with BDI used afterwards. Further, it is insufficient to simply state 'I would use a questionnaire similar to K-SAS' (such as when writing about pyromania, for example).

Answers to **part (a)** questions in this section should include an appropriate plan, have applied a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully, to show good understanding. Candidates should also include appropriate 'general' methodological features such as sample, sampling technique and location of the study. 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 it would be obtained. Elaboration of these general sentences should be included.

In **part (b)(i)**, candidates should describe some relevant psychological knowledge that the whole question is based on. If the question, for example, asks about ways in which pain can be measured, then candidates should describe relevant measures.

In **part (b)(ii)**, candidates should explain what aspects of this psychological knowledge their **part (a)** plan is based on. Answers to these two question parts must be linked.

In **part (c)**, candidates must refer to what they did in their specific plan rather than give a generic answer that could apply to any study. Use of an example or quoting from their plan would be ideal.

Section B can be considered as follows: A teacher teaches a sub-topic from the syllabus and gives the candidate some psychological knowledge. The teacher then tells each candidate to plan a study using method 'x' to investigate some part of that sub-topic. The candidate plans the study using the psychological knowledge of the sub-topic and they use their methodological knowledge about method 'x'. In the examination, **part (a)** is the plan; **part (b)(i)** is the sub-topic knowledge and **part (b)(ii)** is how the knowledge was used to construct the plan. Exam question **parts (c)(i)**, **(ii)** and **(iii)** then ask about some methodological decisions and evaluation about the plan.

Comments specific to questions

Section A

Question 1

(a) (i) Candidates often gave two examples of questions directly from the key study such as: 'they were hostile towards me' and 'they were laughing at me' and were awarded full marks. Other candidates guessed and although sometimes this showed knowledge of the VR-Paranoia questionnaire, such answers could not be awarded marks if they were too vague to be associated with any question.

(ii) The VR Paranoia questionnaire included: 15 items, was a Likert-type scale, was a 4-point scale, had ratings 0 – 3 and ratings ranged from 'do not agree' to 'totally agree'. Any two of these features were awarded marks; any other answer was incorrect.

(b) A number of responses stated a weakness of a participant such as 'they might give socially desirable answers' which could not gain credit as it was not a weakness of the VR-Paranoia questionnaire itself. Correct answers might include the rating scale of 0 – 3 which could be said to be too narrow, or the ambiguity in choosing responses of either 'moderately agree' or 'totally agree'.

(c) Most candidates successfully provided two strengths of using virtual reality and were awarded 2 marks. To be awarded the remaining 2 marks allocated for this question, candidates needed to mention the study by Freeman et al. in their answers (and schizophrenia). Generic answers were awarded partial marks, such as 'VR controls the environment' because they are correct. Full marks can only be awarded if the generic answer is related to the question/the study on which the question is based. For example, when adding (to the generic point) 'allowing a range of situations and settings to be investigated with different types of schizophrenic patients'.

Question 2

(a) Answers to this question were often very detailed (10 lines plus) and often included irrelevant detail, such as her husband's problems. Candidates are advised to write more concise answers for 2-mark questions. Apposite was the treatment process, namely pairing imagery of nausea and vomiting with the act of stealing.

(b) Candidates often wrote about a questionnaire without any detail of format, actual questions or how questions could be answered. There was often no mention of how the effectiveness of covert sensitisation would be tested.

(c) There were many appropriate points made here, with most candidates stating that the main strength is that no drugs are used in psychological treatments. The specific treatment of the question was kleptomania and very few candidates related their point about drugs to kleptomania. There were only a small number of candidates who referred to kleptomania in their answers when this was an essential component of the question.

Question 3

(a) Candidates were required to describe two dependent variables from the Becker et al. key study. Many answers were incorrect despite DV's being an essential component of any key study. Correct answers were awarded 1 mark for identifying 'taste intensity evaluation', 'product evaluation' or 'price expectation'. Further marks could be gained through elaboration of any two of these, such as 'sharp bitter and mild on a 7-point scale' for taste intensity.

(b) This question was answered fully and correctly by many candidates who clearly explained that measuring sensitivity in participants and placing them into two different groups controlled for participant variables.

(c) Similar to other **part (c)** questions asking about weaknesses, many responses to this question included 'social desirability'. This is not a weakness of a seven-point scale. Creditworthy answers included: 'using a 7-point scale provides quantitative data and there is no opportunity to gather qualitative data' which would be awarded 1 mark. To achieve a second mark this general comment needed to be related to the study/question, such as writing 'so there is no opportunity to give a reason for choice in relation to yogurt pot/food package design'.

Question 4

(a) Nearly all candidates were awarded full marks for correctly identifying and outlining a type of interior store design. A 'racetrack/boutique' style was the most common with many candidates naming a store using that type of design in addition to outlining how such a store is laid out.

(b) Candidates commonly suggested conducting a questionnaire or interview outside an actual shop where people would be asked about whether they liked it or not. Only a few candidates showed knowledge and understanding when asking about a specific feature such as the 'free-form arrangement' or about the format and technique of their questionnaire/interview.

(c) Many answers gave two weaknesses of conducting virtual store layout studies, such as the lack of ecological validity, lack of idiographic input and potential side effects of using VR. Answers often did not consider or show knowledge of different store layouts.

Question 5

(a) (i) Candidates could often state that the Bridge et al. study was longitudinal because it was conducted over a period of time. Fewer candidates were able to access a further mark by stating that it lasted for six weeks or that the effectiveness of the programme was assessed using the Leeds General Scale and the Profile of Mood States.

(ii) Many candidates stated that the results for the treatment group were 'better' than the control group. Such answers were too vague to be credited. Answers needed to include specific groups (any two from 'relaxation group', 'relaxation plus imagery group' and controls) and needed to include some numbers to earn the second available mark, such as 61.7 after 6 weeks for the relaxation group compared to 53.9 after 6 weeks for the relaxation plus imagery group.

(b) Suggesting a bar chart with bars for initial/six-week tests and for each of the three groups gained full credit. A line graph with explanation was also creditworthy. Suggesting a histogram, or scatter graph did not receive credit.

(c) Most candidates could provide two strengths of using longitudinal studies, often making the point that it tracks the same individual over a period of time. By adding the comment 'and by using the Leeds scale at different time intervals such as six weeks to assess the effectiveness of the programme' would be an excellent answer and gain full marks. To gain the highest marks, candidates must always relate the strength or weakness to the study/sub-topic of the question.

Question 6

(a) Maximum marks were awarded to candidates writing, for example, 'objective is fact, a measure that cannot be disputed. For example, if a drug is detected in blood or urine then it is an objective fact'. Some candidates could not name a biological measure, often giving pill counting/Trackcap as an example.

(b) The most common suggestion was Trackcap which could receive credit for this question part because the number of pills leaving a dispenser is objective fact. Some candidates wrote about repeat prescriptions which is also creditworthy because a prescription has been collected from a pharmacy that is also objective fact.

(c) There were some very good answers to this question. Candidates needed to relate the weakness of the biological measure to adherence to gain full credit. For example, writing 'collecting blood is invasive and spitting into a salivette may be embarrassing' is correct, but here, in this typical answer, there is no mention of adherence. All **part (c)** answers like this one must be related/linked to the sub-topic of the question in order to receive full marks.

Question 7

(a) (i) The two independent variables for this key study by Cuadrado et al. were correctly identified by nearly all candidates by stating 'the leader's sex (male v female)' and 'the leadership style (male stereotypical v female stereotypical)'. Some candidates incorrectly suggested that the male and female participants was an IV.

(ii) Like **Question 3(a),** candidates should be able to identify the dependent variables of a key study, yet many did not. There were three DV's: the leader's competence; the leader's efficacy/effectiveness and 'evaluation'. All three were measured using a 7-point Likert scale.

(b) Suggestions included the conducting of field experiments, questionnaires, interviews and observations, all of which were acceptable. Answers being awarded full marks explained how their investigation would be conducted and specifically linked their suggestion to women's managerial position in organisations.

(c) As with previous questions, candidates need to elaborate on generic points in order to achieve full marks. Writing 'the study is conducted in an artificial environment' is a weakness of a laboratory

experiment but, without reference to the study in question, the answer is meaningless. Any reference to women in managerial positions would receive credit.

Question 8

(a) Industrial sabotage is defined as 'any behaviour by an employee which is intended to inflict a production or profit loss for the targeted organisation and, commonly, this involves slowdown or destruction of machinery'. Many answers received maximum marks but there were two common errors which could not be awarded marks **(i)** considering sabotage as something not industrial (as this is the organisations option) **(ii)** considering stealing from a co-worker as a form of sabotage.

(b) The question asked for an open question about reasons for sabotage. Amongst excellent answers there were two common errors: **(i)** some candidates did not ask an open question, such as failing to use the words 'describe' or 'explain why'. If a question can be answered with a single word, it is not an open question **(ii)** the question did not ask about a 'reason' for sabotage.

(c) Candidates made perfectly valid points to this question but often did not relate the point to the study in question. Typical weaknesses were workers not being truthful due to 'social desirability', not wanting to be sacked, or exaggerating or underemphasising details. Weaknesses were often relevant but candidates frequently did not refer to sabotage in their answers when it was essential to do so.

Section B

Question 9

(a) Many candidates decided to conduct an experiment to investigate the effectiveness of ECT compared to a different treatment method such as drugs or 'befriending'. This approach did not answer the question set on side effects/problems of ECT, and so no marks could be awarded. Candidates answering the question focused on side effects from treatment of ECT conducted over different time periods which addressed the longitudinal component of the question. For a DV, some candidates correctly focused on testing for side effects such as memory loss; others incorrectly focused on whether symptoms had improved or not.

To maximise marks candidates should **(i)** use the named method and its specific features; **(ii)** apply that method to the question; **(iii)** answer the specific question asked (in this instance, on potential side effects caused by ECT).

(b) (i) For psychological knowledge, some candidates correctly focused on the side effects of ECT, typically memory loss. Others wrote incorrectly about positive and negative symptoms. Those who compared ECT to a different method often wrote exclusively about drugs or befriending, which did not discuss ECT and side effects. This approach was inappropriate and no marks could be awarded. Candidates must describe knowledge of what is in the question that is set.

(ii) Candidates achieving top marks for this question part explained how the memory loss described in **part (b)(i)** was tested in their **part (a)** plan, often it was the DV. Some candidates incorrectly gave strengths and weaknesses of their plan and others wrote answers that did not relate to what was described in **part (b)(i)**. When answering questions in **part (b)(ii)**, the aim is to explain how the psychological knowledge described in **part (b)(i)**, informs the plan of **part (a)**.

(c) (i) Candidates often included a number of controls in their **part (a)** answer and most could give a reason why they used a control. Often, candidates did not give an example from what they had written earlier to support the reason.

(ii) The main weakness given was that participants may drop out of the study. A number of candidates suggested that 'researchers may form a relationship/become attached to the participant'. This is incorrect; to do this would be reason for practitioner dismissal. Because 'attachment' was referred to in a 1954 study (Thigpen and Cleckley), it should not be used as a weakness for any other longitudinal study.

(iii) Quite a few candidates did not provide a hypothesis in **part (a)** and so could not answer this question. Other candidates provided an answer and stated 'because it predicts a direction'. This is insufficient as the answer must be related to this specific plan. A top mark response might include 'I

used a directional hypothesis because it was predicted, on the basis of psychological knowledge described in **(b)(ii)**, that ECT would cause memory loss.'

Question 10

(a) This question required a plan to use an interview with the bullet points of the question guiding the candidate to include interview format and interview technique in their plans. Whilst many did this, others decided to conduct an experiment, describing specific features of their experiment rather than describing the specific features of their interview. This wasted time for candidates and resulted in a lower mark overall. Candidates are advised to base their answer on the method that is named in the question. A number of answers were showed very little methodological or psychological knowledge. Candidates should choose a question that they have studied rather than one that may appeal to their personal knowledge.

(b) (i) For psychological knowledge, the most appropriate research would be that by Mower et al. (2012) who studied window displays. Some candidates impressed with a detailed knowledge of the study, describing the use of a 'female mannequin wearing fashionable clothes' for example. Other candidates wrote very anecdotal answers, often describing what they have seen in shop windows.

(ii) This question part required a link to show how what was described in **part (b)(i)** informed the plan in **part (a)**. For example, in **part (b)(i)** it could have been written that Mower et al. used mannequins. Then, in **part (b)(ii)** it could be explained that mannequins were also used or explained that a modified display was used instead.

(c) (i) There was a common assumption that interview format is concerned with using open and closed questions which is not correct. Interview format is whether the format is structured, semi-structured or unstructured. There is also the common assumption that a structured interview is automatically closed questions and that an unstructured interview is open questions. This is also incorrect. A structured interview is where all participants receive the same questions in the same order whether they are open or closed.

(ii) Candidates who answered **part (c)(i)** incorrectly often answered this question part incorrectly as well. There were some candidates who answered both question parts correctly and applied their reason/weakness to their plan and were awarded full marks.

(iii) Most candidates answered this question part incorrectly because they confused interview technique and interview format. For clarification, interview technique is whether the interview is by telephone or face-to-face. Further, most candidates used a face-to-face interview in their **part (a)** plan but were unable to give a reason why they had used that technique here.

Question 11

(a) There were many excellent plans which investigated the question directly. Weaker responses were characterised by: **(i)** not addressing the question of reasons why people fail to attend appointments. Some candidates looked at rewards for immunising children instead. **(ii)** not designing an appropriate postal questionnaire, some candidates suggested interviewing participants. The comments at the end of **Question 9(a)** should be noted.

(b) (i) For psychological knowledge, the most appropriate research was that of Laba et al. (2012) which looked at rational non-adherence, or the health belief model which provides a number of reasons for non-adherence. Some candidates wrote about the Yokley and Glenwick study on immunisation but focused incorrectly on encouraging people to attend clinics rather than the reasons for failure to attend as the question required.

(ii) Candidates describing reasons for adherence in **part (b)(i)** sometimes explained how they had used that knowledge in their plan. Many responses did not include this. Some wrote more detail, extending their **part (b)(i)** answer and others evaluated their plan. Both these latter types of answer could not be awarded marks.

(c) (i) Question format is about whether questions are open and/or closed. Candidates could choose open, or closed, or a combination of the two and here the reason for that choice is explained. Often candidates gave a generic reason, such as 'so I could gather quantitative data', without any reference to how this would inform the question of reasons for failure to attend an appointment.

(ii) Sending an email (as was done in many plans) is not a postal questionnaire and so some candidates were unable to score marks here. The most impressive answers explained that sending a postal questionnaire was of little use because those failing to attend appointments would also fail to return postal questionnaires.

(iii) Many candidates chose an opportunity sampling technique or a volunteer sample without considering how this would apply to sending out postal questionnaires. A few candidates chose a random sample without explaining how the sample would be random.

Question 12

(a) Many answers did not address the named method so were unable to gain full credit. The focus was structured observation so answers planning laboratory experiments or questionnaires could not be awarded marks. If the method is an observation, it is expected that candidates will include all the specific features of observations (see mark scheme for details). Some candidates did not seem to know about sabotage, specifically that of machinery. Some candidates wrote superb answers addressing all the relevant feature needed.

(b) (i) Relevant psychological knowledge for this question was the key study on reasons for sabotage in the workplace by Giacalone and Rosenfeld (1987). They list forms and reasons for sabotage. Their 29 'forms' include creating work slowdowns and 'forgetting' to turn on/off a switch. These or any of the others could be used as a basis for an observation.

(ii) Candidates describing a form of sabotage in **part (b)(i)** often designed behavioural categories to investigate some aspect of it. For example, 'creating work slowdowns' could have the categories of 'number of times within a week'; 'the duration of the slowdown'; 'the cause of the slowdown' etc. In the strongest answers there was a clear link shown between the description in **part (b)(i)** and the plan in **part (a)**.

(c) (i) Because some candidates did not have any behavioural categories in their **part (a)** answer, they could not answer this question. A few candidates in this position decided to go back to **part (a)** and add some categories, which is acceptable. All candidates are strongly advised to read the question fully and follow the instructions. In this instance, 'your plan must include details about behavioural categories'.

(ii) Candidates with no categories also struggled to answer this question. Some appropriate weaknesses were included by many candidates, often referring to sabotage behaviours that were observed that were not in their original list of behaviours to be observed.

(iii) The question 'steps for making the study reliable' appeared to confuse some candidates. Some candidates wrote about the procedure of the study, standardising what they could. The strongest answers simply referred to their plan where inter-rater reliability could be checked to see how similar/consistent their recordings of sabotage behaviour were.

PSYCHOLOGY

Paper 9990/42

Specialist Options: Application and Research Methods

Key messages

- What has been learned from the AS component of the syllabus should be transferred to the A2 component. For example, at AS candidates learn about methodology, such as experiments, which also apply to A2.
- Questions should be read carefully ensuring that the focus is on what the question asks.
- For ***Section A*** answers, candidates should relate their answer to the study in question or include an example. Questions frequently end with 'in this study' and so the answer should be related to that specific topic area/study.
- All terminology should be explained. Writing 'it is valid and reliable' for example, is insufficient without explanation, application or example.
- The syllabus includes for 'example studies' such as 'e.g., Oldham and Brass (1979)'. Example studies can be substituted for alternatives, but these alternatives must cover the same or very similar content to the example study. If the Oldham and Brass study is substituted, the alternative study must be about a move to open plan offices and the data that was gathered from that move. The alternative cannot be about something different.

General comments

Some candidates answered questions from one option only. Other candidates, who correctly answered two options, sometimes performed considerably better in one option than the other.

Many candidates answered two questions from ***Section B*** instead of one (only one of these ***Section B*** responses can receive credit). Candidates are advised to read the instructions on the front cover of the question paper and to read the heading instructions for each question section.

Candidates should double check that the terminology they use in their answers is correct. Often terms such as reliability and validity were used interchangeably, as were qualitative and quantitative, and independent and dependent variables. There was also confusion with the terms format and technique in relation to questionnaires and interviews.

Section A

Question **part (c)** requires a general evaluative point that could relate to any study (such as a strength or weakness of a method) but it also requires for the general point to be related to the specific sub-topic/study in the question. Answers often included strengths and weaknesses but these were often not related to the question, and so restricted marks.

Candidates should not use psychological terms without explanation. Frequently answers were limited to 'it is reductionist' or 'it is useful in everyday life' without further explanation. Stating 'it is reductionist' does not automatically identify it as a strength or weakness.

Candidates should not use the terms reliability and validity to answer every **part (c)** question for three reasons: **(i)** they do not apply to most questions and so cannot be awarded marks, **(ii)** candidates using the terms often do not know how they apply to the specific question and **(iii)** candidates often confuse the terms.

Section B

Candidates should only answer one question from this Section.

Many candidates appeared to assume that they must conduct an experiment whatever the question. An interview, questionnaire or observation are methods independent of an experiment and candidates should not try to make other methods 'fit' into an experimental format.

Some candidates evaluate their plan in **part (a)** by listing strengths and weaknesses. This should not be done because: the question does not ask for evaluation; there are no AO3 marks allocated to evaluation; evaluation is done in **Questions (c)(i)**, **(c)(ii)** and **(c)(iii)**.

Some candidates included a paragraph of results. This achieves no marks because the question asks for a plan only. Further, the proposed plan has not been carried out so no actual results are gathered.

Candidates need to know the distinction between questionnaire format and technique, and interview format and technique, as stated on the syllabus: Questionnaire technique: paper and pencil (i.e., done by a person with the researcher present), online or postal. Questionnaire format: open and/or closed questions. Interview technique: telephone or face-to-face. Interview format: structured, semi-structured, unstructured.

When using psychometric tests candidates should not use acronyms unless the full title of it is provided first. For example, 'Beck Depression Inventory (BDI)' is fine, with BDI used afterwards. Further, it is insufficient to simply state 'I would use a questionnaire similar to K-SAS' (such as when writing about pyromania, for example).

Answers to **part (a)** questions in this section should include an appropriate plan, have applied a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully, to show good understanding. Candidates should also include appropriate 'general' methodological features such as sample, sampling technique and location of the study. 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 it would be obtained. Elaboration of these general sentences should be included.

In **part (b)(i)**, candidates should describe some relevant psychological knowledge that the whole question is based on. If the question, for example, asks about ways in which pain can be measured, then candidates should describe relevant measures.

In **part (b)(ii)**, candidates should explain what aspects of this psychological knowledge their **part (a)** plan is based on. These two question parts must be linked.

In **part (c)**, candidates must refer to what they did in their specific plan rather than give a generic answer that could apply to any study. Use of an example or quoting from their plan would be ideal.

Section B can be considered as follows: A teacher teaches a sub-topic from the syllabus and gives the candidate some psychological knowledge. The teacher then tells each candidate to plan a study using method 'x' to investigate some part of that sub-topic. The candidate plans the study using the psychological knowledge of the sub-topic and they use their methodological knowledge about method 'x'. In the examination, **part (a)** is the plan; **part (b)(i)** is the sub-topic knowledge and **part (b)(ii)** is how the knowledge was used to construct the plan. Exam question **parts (c)(i)**, **(ii)** and **(iii)** then ask about some methodological decisions and evaluation about the plan.

Comments specific to questions

Section A

Question 1

(a) (i) Candidates often gave two criteria (for inclusion) directly from the key study such as: diagnosis of OCD, score of 16 or more on the Y-BOCS, and were aged between 16 and 65 years and all of these earned credit. Other candidates guessed at criteria and although sometimes this showed knowledge of OCD, such answers could not be awarded marks if they were not criteria used specifically by Lovell et al. (2006).

(ii) Like **Question 1(a)(i),** reasons (for exclusion) needed to be those stated by Lovell et al. These included: obsessional slowness, organic brain disease, a diagnosis of substance misuse, severe

depression or with suicidal intent or patients who had been on anti-depressants or anxiolytics for less than 3 months. Other answers could not be awarded marks.

(b) The key study by Lovell et al. did not use a control group because it was not needed; the aim was to see if CBT by telephone was as good as face-to-face. The study was not testing the effectiveness of CBT where a control group would be needed.

(c) Most candidates had little difficulty in providing two strengths of excluding participants from studies and generic answers like these were awarded partial marks, because they were correct. Full marks could only be awarded if the generic strength (or weakness) was related to the question/study, in this instance studies of OCD.

Question 2

(a) Answers to this question needed to include two (of three) important aspects for full marks: **(i)** the words 'during the past week', **(ii)** one actual question from the K-SAS, such as 'How many times did you steal?' or words very similar and **(iii)** the scale used to assess a specific question because a different scale was often used for different questions. 'A four-point scale' was too vague.

(b) Candidates were required to suggest how the reliability of the K-SAS could be tested. Many candidates wrote about test-retest reliability (same person on different occasions) and were awarded 2 marks. Use of split-half was also creditworthy. Many candidates incorrectly wrote about validity.

(c) Many candidates incorrectly wrote about social desirability which is a weakness of a participant doing any study/questionnaire; it is not specifically a weakness of the K-SAS. Many candidates wrote correctly about the different rating scales for different questions, the overlap of rating scales and the use of the ambiguous word 'approximately'.

Question 3

(a) The question asked about the procedure of study 3 specifically, so any answer that was not about study 3 or about the procedure could not be awarded marks. Details about the telephone interview, the two types of message (image or quality) and the willingness to try the shampoo, for example, were all creditworthy.

(b) There were only two ethical guidelines broken in this study which were deception and lack of informed consent. Candidates suggesting alternatives, such as harm, scored 0 marks. Deception, for example, was involved when researchers **(i)** gave a false name, **(ii)** claimed that they were working for a marketing research firm, **(iii)** claimed they were researching a shampoo.

(c) Like other **part (c)** questions, candidates often gave a strength and weakness without relating it to the question. Those addressing the question specifically, were often awarded full marks. These responses often included an example about the two images, the lack of non-verbal communication or the belief that it was genuinely market research.

Question 4

(a) A small number of candidates confused eye tracking with eye-magnets. For 1 mark, candidates needed to mention the preference for a centrally located item and for a second mark some aspect from the 'one study' they were outlining. From the Atalay et al. study, this could be about the use of a planogram, product categories (vitamins, etc.) or choosing the central item even when it is not placed in the centre of the shelf or the visual field.

(b) Suggestions included the conducting of field experiments, laboratory experiments, questionnaires, interviews and observations (with or without CCTV). Answers being awarded full marks explained how their investigation would be conducted, for example by simply asking participants where they first look when looking at a product on a shelf. Some candidates suggested using EEG or fMRI, but neither of these would replace eye movement.

(c) Many answers gave a strength and a weakness of conducting a study using eye-tracking equipment such as the device being unnatural or uncomfortable. Many answers were not related to

the Atalay et al. study (or alternative study) on product choice or even consumer behaviour and so could not be awarded the 2 additional marks for application.

Question 5

(a) (i) Most candidates were awarded 2 marks for correctly identifying **(i)** the Wong-Baker Faces Pain Rating Scale which has 6 faces showing 0 – 10 increasing levels of pain and **(ii)** the Faces Pain Scale (FPS-R) which also has 6 faces with a 0 – 10 scale (0, 2, 4, 6, 8, 10).

(ii) Marks were awarded for outlining **(i)** the Visual Analogue Scale (VAS), a line drawn on paper with a severity scale and **(ii)** the Coloured Analogue Scale (CAS) which has a line on a piece of paper with pain severity ranging from 0/green to 100/red).

(b) Some candidates suggested using a rating scale which was incorrect given the wording of the question. Others suggested using the UAB but this also uses a rating scale. Correct answers were those suggesting observation of behaviours such as crying, grimacing, etc. Also creditworthy was a clinical interview where the child could be asked questions about their pain.

(c) The question specifically asked about weaknesses of the rating scales used by the 3 – 8-year-olds. Answers that were about social desirability, alternative pain measures or the rating scales used by the 9 – 15-year-olds could not be awarded marks. Correct answers focused on the narrow range of faces or the lack of understanding by a child in pain rating faces.

Question 6

(a) The syllabus mentions two studies in this sub-topic which gathered salivary cortisol, those by Wang et al. (2005) and Evans and Wener (2007). A brief outline was often awarded 2 marks. These are 'e.g.' studies so alternatives could also be credited. If an alternative is used, then sufficient detail needs to be included to allow an Examiner to look up the study to test its appropriateness.

(b) Most candidates were awarded at least 1 mark for suggesting how saliva can be collected from a child. Most suggestions involved playing a game or giving some form of positive reinforcement.

(c) There were some excellent answers focusing on the strengths of biological measures such as salivary cortisol and blood and urine tests. There were many answers that ignored the 'biological measure' component of the question e.g. candidates instead writing about pill counting. A number of candidates wrote incorrectly about the strengths of children using the 'funhaler'.

Question 7

(a) The key study by Landry et al. included two independent variables: autonomy supporting and autonomy controlling/threatening. Nearly all candidates were awarded 2 marks for identifying these variables and many candidates provided some explanation of what the terms meant and were awarded 2 further marks for elaboration.

(b) Many candidates correctly wrote about the validity of the results (if the manipulation check had not been done) and top answers related this to the Landry et al. study. Other candidates wrote about validity but did not relate this to Landry et al. and were awarded 1 mark. Other candidates incorrectly wrote about reliability which was not applicable.

(c) There were three types of answer: **(i)** those who knew the design was independent, gave a strength and weakness of it and applied these to the Landry et al. study, **(ii)** those who knew the design but did not apply it and **(iii)** those who did not know what an experimental design was, often giving a strength and weakness of laboratory experiments.

Question 8

(a) Answers to this question required an *outline* and so *identification* of the styles could only be awarded 1 mark. A full mark answer for example might be 'a leader makes unilateral decisions, which will be carried out by workers, subordinate to them, and who have no involvement in the decision-making process'.

(b) 2 marks were awarded to answers including a closed question (rather than a statement) which was **(i)** related to the directive autocrat style and **(ii)** had a response choice of either yes/no or some rating scale. There were many partial mark answers, where one or more of the above features were absent, and there were a few open questions which were incorrect.

(c) A number of candidates did not know the term structured observation and some substituted it for a covert observation. Those using structured observation gave strengths but often did not apply it to a directive autocrat style. Some candidates applied all the required components and were awarded full marks.

Section B

Question 9

(a) Many candidates decided to conduct an experiment and wasted time writing about IV, DV and other aspects of experiments. The required method was a *questionnaire* so the detailed focus should have been on that, mentioning the technique (paper/pencil, online, postal), format (open and/or closed), examples of questions, question scoring/interpretation and the number of questions. A large number of candidates did not address these features; they simply stated 'a questionnaire similar to K-SAS', or even used a K-SAS, which did not show knowledge or understanding of questionnaire design or pyromania. A second weakness was that some candidates did not seem to know about imaginal desensitisation other than the words themselves, and others confused it with covert sensitisation. To maximise marks, candidates should **(i)** use the named method and its specific features as outlined above, **(ii)** apply that method to the question (in this instance showing how imaginal desensitisation is used to treat pyromania) and **(iii)** answer the question (in this instance address how *effectiveness* can be determined).

(b) (i) For psychological knowledge, most candidates referred to diagnostic criteria for pyromania or explained the treatment of imaginal desensitisation e.g. that by Blaszczynski and Nower (2003). This involves progressive muscle relaxation and then the person visualises themselves being exposed to the situation that triggers the drive to carry out the impulsive behaviour. It does not involve being shown images or vomiting as some candidates believed.

(ii) Candidates usually used the procedure outlined by Blaszczynski and Nower. When answering questions like this, it might be helpful for candidates to begin their answer with 'my plan involved doing 'x' because…'

(c) (i) Most candidates answered this question part incorrectly. The question stated questionnaire *technique* and **not** questionnaire *format*. Technique relates to whether the questionnaire is given to participants face-to-face (i.e. 'paper and pencil'), postal or online. Writing about an open or closed question format was incorrect.

(ii) Many candidates writing about question format in **Question 9(c)(i)** also provided incorrect answers to this question part when giving a weakness of their question format. Candidates answering the question correctly often gave a correct weakness but did not relate their answer to their plan.

(iii) Nearly all candidates were awarded 1 mark for giving a general reason for using a rating scale. Most candidates could not be awarded a second mark because they did not give a reason for the choice of their specific rating scale and did not give an example from their plan to support their answer.

Question 10

(a) This question required use of a face-to-face interview, which many candidates did not do, deciding to conduct an experiment instead. This is acceptable but the detail must be on the named method (for which marks will be awarded) not an alternative method. Many candidates knew the disrupt-then-reframe technique but often became confused when introducing additional IV's such as reframe only or when using the technique itself. The specific features (see mark scheme) of interviews were often absent. Like all questions in ***Section B*** to maximise marks candidates should **(i)** use the named method and its specific features, **(ii)** apply that method to the question and **(iii)** answer the question (in this instance address how *effectiveness* of DTR can be determined).

(b) (i) For psychological knowledge, the most appropriate research would be that by Kardes et al. (2007) who found that confusing consumers with disruptive/confusing information and then re-wording (reframing) in a much clearer way resulted in increased sales. For example, 'The price is now 100 Eurocents (approximately 100 pennies) that's 1 Euro. It's a bargain!'

(ii) This question part required an explanation to show how what was described in **Question 10(b)(i)** informed the plan in **part (a)**. For example, in **part (b)(i)** the DTR technique and example was described. In **part (b)(ii)** it could be explained why the example of '100 eurocents and 1 Euro' was used or explained why a modified version of this was used instead.

(c) (i) Question format is about whether questions are open and/or closed. Candidates could choose open, or closed or a combination of the two and here the reason for that choice is explained. Most candidates opted to use closed questions but gave a generic answer, such as 'so I could gather quantitative data', and frequently did not mention what the quantitative data was or how it could determine effectiveness.

(ii) For candidates choosing open questions the weakness was nearly always 'no quantitative data' and for those choosing closed questions 'no qualitative data'. Very few candidates provided any elaboration and very few linked this to what they had done in their plan; both essential aspects for full marks.

(iii) Many candidates chose an opportunity sampling technique and provided an appropriate reason for this choice; others chose a volunteer sample and, again, often explained why this was appropriate to their plan. A few candidates chose a random sample which was incorrect because a random sample is where everyone in a population has an equal chance of participating. Choosing people walking into a store (or whatever location chosen for the study) is not a random sample; it is an opportunity sample.

Question 11

(a) There were many excellent plans provided by candidates. There were also many which were unable to achieve more than bottom band marks because candidates did not know what stress inoculation was. Stress inoculation (in the term itself) is something to try to *prevent* stress; it is not a stress reduction technique to help stressed people manage stress. Answers achieving high marks understood the inoculation process and included conceptualisation, acquisition and application. Crucially, top answers also assessed effectiveness using a semi-structured interview, with answers at the bottom end of the mark range not knowing what a semi-structured interview was. Note comments about maximising marks made at the end of **Question 9** and **Question 10**.

(b) (i) For psychological knowledge, the most appropriate research was that of Meichenbaum (1985) who prevented stress using self-instructional training and inoculation therapy. This technique was often described in appropriate detail. Candidates writing about the Bridge et al. study were awarded 0 marks because that study was about managing stress, not stress inoculation.

(ii) Candidates describing stress inoculation in **part (b)(i)** explained how they had used that knowledge, such as the three stages of conceptualisation, acquisition and application in their plan. Candidates who did not know anything about stress inoculation often wrote nothing at all.

(c) (i) Most candidates answered this question part incorrectly because they confused interview *technique* and interview *format*. In this instance, the question asked about interview technique which is whether the interview is by telephone or face-to-face. Candidates writing about interview format: structured, semi-structured, unstructured were awarded no marks. Those writing about interview technique often provided a good reason for their choice of either face-to-face or telephone interview and related their answer to stress inoculation.

(ii) Many candidates did not seem to know the meaning of the term semi-structured (interview), and so could not give a correct strength. Many candidates assumed incorrectly that a structured interview gathers quantitative data and an unstructured interview gathers qualitative data. All methodological knowledge from AS should be carried forward and be able to be applied at A2.

(iii) Like many other **part (c)** answers, responses here were often single, generic statements such as 'I chose to use quantitative data'. This type of answer is not a *reason* for choosing this type of data; and there is no mention at all of how this type of data was used in the **part (a)** plan. An example of

a 2 mark answer might be 'I chose quantitative data because the rating out of ten I used (which appears in **part (a)**) can be used to determine the effectiveness of stress inoculation compared to 'x'' (if an alternative technique or no technique was used in **part (a)**).

Question 12

(a) Many responses did not answer the question set. The focus of this question was on types of achievement motivation (see **part (b)(i)** below) and not intrinsic and extrinsic motivation. Questionnaires were often worded appropriately with candidates choosing closed questions, open questions and sometimes a combination of both. The term 'online' questionnaire saw some candidates use a 'world-wide' online sample whilst others had participants using a computer in a laboratory. As has been mentioned, to maximise marks in ***Section B*** candidates should **(i)** use the named method and its specific features, **(ii)** apply that method to the question and **(iii)** answer the question (in this instance address type of achievement motivation).

(b) (i) Relevant psychological knowledge here was McClelland's achievement motivation theory (1961) which suggests that there are three work-related needs: need for achievement, need for affiliation and need for power. Candidates describing either Maslow's hierarchy, or the difference between extrinsic and intrinsic motivation were awarded no marks.

(ii) Candidates describing McClelland's three needs in **part (b)(i)** often designed questions based on the three needs and top answers explained how a specific question applied to a specific need. Candidates using Maslow or extrinsic and intrinsic motivation planned incorrect questions and so could not be awarded any marks.

(c) (i) This question asked about the *sample of participants* and not the *sampling technique.* Candidates opting for the latter were awarded 0 marks. Top answers mentioned reasons for their choice of sample of participants such as the number, sex/gender balance, location, age range and specific features such as the organisation they worked in.

(ii) The weaknesses included were dependent on the features mentioned in **part (c)(i)** such as the restricted age range or the inability to generalise from participants from one organisation.

(iii) Question format is about whether questions are open and/or closed and this should have been specifically addressed in **part (a)** because it is part of the question itself (the bullet points). In this question part, the reason for that choice should be explained. Most candidates opted for use of a closed question but often because the wrong questions were asked (incorrect focus on Maslow or intrinsic/extrinsic motivation) so no credit could be awarded.

PSYCHOLOGY

Paper 9990/43

Specialist Options: Applications and Research Methods

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