PSYCHOLOGY

Paper 9990/11
Approaches, issues and debates

Key messages

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

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

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

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

Candidates need to be able to explain similarities and/or differences between studies based on psychology**.** Brief, common-sense responses can rarely be credited but these were presented some candidates.

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

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

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

General comments

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

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

There were some blank responses. Candidates are encouraged to attempt all questions even if they are unsure of the answer they are providing, as they may be able to make some creditable points.

Comments on specific questions

Question 1

(a) A majority of responses correctly stated the number of participants studied intensively by Dement and Kleitman. Common incorrect responses included the total number of participants used in the study, or how many were used as additional data.

(b) Stronger responses could clearly outline the part of the procedure linked to investigating dream recall. Common points made by candidates included being woken by a doorbell sound, recording the dream into a tape recorder, and being woken up in REM or NREM sleep. Common incorrect responses included estimation of dream duration. It is important for candidates to read the question carefully to ensure that they are providing the correct part of the procedure.

Question 2

(a) There were few correct responses to this question. Some candidates could name two stimuli from the correct part of the hierarchy. However, a majority of responses were incorrect and some included generic descriptions of buttons or scenarios not part of the hierarchy which did not answer the question set. There were some blank responses to this question.

(b) There were some clear, concise responses to this question. For example, highlighting the role of disgust in the development and the maintenance of a phobia. However, there were many responses that provided a direct result from the case study. Conclusions are generic descriptions of the outcome of a study, whereas results are factual data (quantitative and/or qualitative) provided directly from the participants. Responses that focused on the boy's button phobia could not be awarded credit here.

Question 3

(a) There were many correct responses to this question. Popular descriptions included the sample size, gender, where they were recruited from, and that they had been diagnosed with AS/HFA. However, there were a significant minority of responses that could not describe three features/ characteristics of the sample or confused the sample with that from one of the other three groups. It is important for candidates to note the number of marks assigned to a question as this typically represents (in short answer questions) the number of correct elements that need to feature in a response.

(b) A minority of responses could state the solution used by Baron-Cohen et al. to the two problems presented in the question. More candidates could note that 36 pairs of eyes were used for the first problem compared to being able to use a glossary for the second problem. Common incorrect responses were using male and female eyes for problem 1 or using simpler words for problem 2.

Question 4

(a) Stronger responses could clearly describe four elements of the Memory or Belief? questionnaire. Popular elements included it being based around three items from a previous questionnaire and that asparagus was a focal item. Incorrect responses tended to describe the function of the questionnaire rather than the actual questionnaire. Some responses simply outlined another questionnaire used in the study by Laney et al. There were some blank responses to this question.

(b) There were some clear, concise responses to this question. For example, noting the potential issue of untruthful responses, or not being able to check if a memory or a belief had actually formed. Other popular choices included the restricted choices available or social desirability reducing validity. There was a number of responses that did not provide an example from the study to elaborate on a chosen weakness. Candidates need to be aware that 'in this study' in the question requires an explicit example from that study to be able to access all available marks. Some responses provided a weakness of the study by Laney et al. rather than of the specific questionnaire.

Question 5

(a) A minority of responses could describe a result about helping in the no model condition. The most popular was that more help was given to the 'ill' victim compared to the 'drunk' victim. With results, where possible, there must be a meaningful comparison to be able to access the marks available. Therefore, simply stating more help was given in the 'ill' victim condition is not enough to be awarded the two available marks for a meaningful comparison. Several incorrect responses chose a different result, or a result based on overall trends within the study. In the majority of responses, data was not provided as required in the question. It is important for candidates to read the question carefully. This question had the most blank responses on this paper.

(b) There were some clear, concise responses to this question. For example, noting that the study had high mundane realism, or that it could be replicated as the procedure was standardised. Candidates need to be aware that 'in this study' in the question requires an explicit example from that study to be able to access all available marks. Some responses provided a strength of the study by Piliavin et al. without any example from the study itself to tell the Examiner why it was a strength.

Question 6

Stronger responses could clearly outline the nature versus nurture debate and provide clear examples from the study by Schachter and Singer. Popular examples included natural reactivity to epinephrine (nature) and reacting to the behaviour of the stooge(s) (nurture). There were a significant number of responses that were tautological and could not access marks. For example, stating that the nurture side of the debate is about nurturing cannot be credited as it is does not show understanding. To improve, candidates need to have examples from each Core Study that appropriately support each of the issues and debates at AS Level.

Question 7

Reponses to this question were very varied. Stronger responses gave clear advice about the use of highly arousing imagery with an example. However, many responses focused on explaining why the advice had been given, which was not the focus of the question. The scenario set up a novel situation and told candidates about the study by Canli et al. The question asked for an outline of advice, not to explain the reason for the advice.

Question 8

(a) The minority of responses could clearly outline what was meant by privacy in relation to ethical guidelines. Popular outlines included participants having the right to not reveal personal details or be observed in 'private' locations. However, the majority of responses gave an answer about confidentiality, rather than privacy. It is important that candidates know all ethical guidelines that psychologists must consider when designing and implementing studies.

(b) There were some strong responses provided here that fully engaged with the stimulus material. Don was the most popular choice for the debate by a significant margin. Popular arguments supporting Don included little protection from psychological harm and lack of informed consent. Popular arguments supporting Adria tended to focus on confidentiality. Some responses simply described aspects of ethical guidelines without engaging in the Adria-Don debate so could only gain partial credit. To improve, candidates need to clearly choose one side of the argument and then explain why they support it using evidence from the study. Also, candidates need to know that when outlining ethical guidelines, they need to know the correct terminology (e.g. stating 'protection from psychological harm', rather than just 'protection') and what was explicitly mentioned in the study e.g. there was no informed consent from parents, but a significant minority of responses claimed this had happened.

Question 9

(a) For these types of questions, responses should focus on the general psychology that is being investigated in the study rather than a specific aim of the study. Aspects of the study by Milgram that could gain credit here included a definition of obedience, situational hypotheses, and ideas linked to agency theory. Credit could be given to generic descriptions of the principles for all of these. However, many responses focused too narrowly on the aims of Milgram and what happened in the study.

(b) Stronger responses could clearly explain one similarity and one difference. Popular choices to compare the studies included the ethics, the sample, the collection of quantitative data and the use of a stooge. To improve responses to this type of question, candidates need to choose comparison points that can be developed, using examples from both studies to explain the similarity and/or difference. For example, explaining the different samples can focus on the characteristics and/or sampling technique and allows for a detailed response with a statement about generalisability. However, stating that each study had a different aim does not allow the response to be detailed. There was a number of responses that compared Milgram to a study that was not from the social approach.

Question 10

The strongest responses evaluated the study by Andrade in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of reliability. Common choices included generalisability, ecological validity, ethics, and quantitative data. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Andrade to explicitly support their point. These answers tended to score Level 4 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses, all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by Andrade as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. Some responses appeared to be prepared essays for Andrade without one of the points being about reliability. A response limited to one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did use reliability in an evaluative sense but did not fully explain why it could be a strength and/or a weakness. Several responses did not cover the named issue. Some responses were attempting to focus on real world application which tended to only be awarded partial credit as this question is evaluative in nature and not application. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, real world application is not required. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples from the study to show clear understanding. These are the requirements for a Level 4 response.

PSYCHOLOGY

Paper 9990/12
Approaches, issues and debates

Key messages

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

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

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

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

Candidates need to be able to explain similarities and/or differences between studies based on psychology**.** Brief, common-sense responses can rarely be credited but these are presented by some candidates.

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

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

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

General comments

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

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

There were some blank responses. Candidates are encouraged to attempt all questions even if they are unsure of the answer they are providing, as they may be able to make some creditable points.

Comments on specific questions

Question 1

(a) A majority of responses correctly stated the third label used by Alex the parrot in the study by Pepperberg. Common incorrect responses included the word texture or an example of a shape used (e.g. blue square).

(b) Stronger responses could clearly outline the part of the procedure linked to investigating dream recall. Common points made by candidates included one trainer presenting objects to a different human, being asked questions about the object, and being rewarded for a correct response. Common incorrect responses included giving Alex the reward or punishment, or the procedure that involved the Principal Trainer. It is important for candidates to read the question carefully to ensure that they are providing the correct part of the procedure.

Question 2

(a) The majority of responses could outline the aim of the study Dement and Kleitman. Popular choices included to investigate dream recall differences in REM and NREM, and eye movement patterns linking to dream content. There were some responses that only gave a very brief aim which could only be awarded partial credit. Some responses presented a finding from the study so could not gain any credit.

(b) A minority of responses could outline the context of this type of dream. Many responses referred to dreams that involved driving a car or added a car to an existing dream report (e.g. throwing tomatoes in a car). The minority of responses that did gain credit tended to focus on seeing a man to the left at a junction. There were some blank responses to this question.

Question 3

(a) There were many correct responses to this question. Popular descriptions included the sample size, where they were recruited from, and the sampling technique. However, there was a significant minority of responses that could not describe three features/characteristics of the sample or confused the sample with that from one of the other three groups. It is important for candidates to note the number of marks assigned to a question as this typically represents (in short answer questions) the number of correct elements that need to feature in a response.

(b) A majority of responses could state the solution used by Baron-Cohen et al. to the two problems presented in the question. More candidates could note that four options were used for the first problem compared to being able to use a glossary for the second problem. Common incorrect responses were using male and female eyes for problem 1 or using simpler words for problem 2.

Question 4

(a) Stronger responses could clearly describe four elements of the Food Costs Questionnaire. Popular elements included how much they were willing to pay for items, choices were given including a 'would not buy', and examples of price points. Incorrect responses tended to describe the function of the questionnaire rather than the actual questionnaire. There were a significant minority of responses that described the Restaurant Questionnaire in error.

(b) There were some clear, concise responses to this question. For example, noting the potential issue of responses not matching actual behaviour, and a lack of qualitative data to explore choices. Other popular choices included the restricted choices available or social desirability reducing validity. There was a number of responses that did not provide an example from the study to elaborate on a chosen weakness. Candidates need to be aware that 'in this study' in the question requires an explicit example from that study to be able to access all available marks. Some responses provided a weakness of the study by Laney et al. rather than of the specific questionnaire.

Question 5

(a) A minority of responses could describe a result about the Activity Index from the two groups highlighted in the question. The most popular was the placebo group scoring higher than the other group. With this type of question where the groups are already given, a description then requires a comparison and a description of what a higher/lower score indicated. Several incorrect responses chose a different result, or a result from a different part of the study. In the majority of responses, data was not provided as required in the question. It is important for candidates to read the question carefully. There were some blank responses to this question.

(b) There were some clear, concise responses to this question. For example, noting that the study could be replicated as the procedure was standardised. Candidates need to be aware that 'in this study' in the question requires an explicit example from that study to be able to access all available marks. Some responses provided a strength of the study by Schachter and Singer without any example from the study itself to tell the Examiner why it was a strength.

Question 6

Stronger responses could clearly outline the individual-situational debate and provide clear examples from the study by Milgram. Popular examples included participants stopping at different levels of voltage (individual) and the prods keeping participants on task (situational). There was a significant number of responses that were tautological and could not access marks. For example, stating that the situational side of the debate is about the situation cannot be credited as it does not show understanding. To improve, candidates need to have examples from each Core Study that appropriately support each of the issues and debates at AS-Level. There were some blank responses to this question.

Question 7

Reponses to this question were very varied. Stronger responses gave clear advice about the use of non-aggressive models, or Omar being a good role model. However, many responses focused on explaining why the advice had been given, which was not the focus of the question. The scenario set up a novel situation and told candidates about the study by Bandura et al. The question asked for an outline of advice, not to explain the reason for the advice.

Question 8

(a) The majority of responses could name two features of the sample used in the study by Piliavin et al. Popular choices included the sample size, the sampling techniques used and the location. Common errors included giving features of the victims or models rather than the sample of participants.

(b) There were some strong responses provided here that fully engaged with the stimulus material. Amina was the most popular choice for the debate. Popular arguments supporting Amina included only one location in one city, only victims of two races used, and only male victims used. Some responses simply described aspects of generalisability without engaging in the Saad-Amina debate so could only gain partial credit. To improve, candidates need to clearly choose one side of the argument and then explain why they support it using evidence from the study. Also, candidates need to know that when focusing on generalisability, they must provide reasons why aspects of the study can or cannot be generalised, rather than simply describing the characteristics of the sample.

Question 9

(a) For these types of questions, responses should focus on the general psychology that is being investigated in the study rather than a specific aim of the study. Aspects of the study by Andrade that could gain credit here included dual processing, the role of boredom in attention, and factors affecting concentration. Credit could be given to generic descriptions of the principles for all of these. However, many responses focused too narrowly on the aims of Andrade and what happened in the study.

(b) Stronger responses could clearly explain one similarity and one difference. Popular choices to compare the studies on included ethics, the sample, laboratory studies, and experimental designs. To improve responses to this type of question, candidates need to choose comparison points that can be developed, using examples from both studies to explain the similarity and/or difference. For

example, explaining the different samples can focus on the characteristics and/or sampling technique and allows for a detailed response with a statement about generalisability. However, stating that each study had a different aim does not allow the response to be detailed. There was a number of responses that compared Andrade to a study that was not from the cognitive approach.

Question 10

The strongest responses evaluated the study by Canli 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 validity. Common choices included generalisability, ecological validity, ethics, and quantitative data. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Canli et al. to explicitly support their point. These answers tended to score Level 4 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses, all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by Canli et al. as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. Some responses appeared to be prepared essays for Canli et al. without one of the points being about reliability. A response limited to one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context, which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did use reliability in an evaluative sense but did not fully explain why it could be a strength and/or a weakness. Several responses did not cover the named issue. Some responses were attempting to focus on real world application which tended to only be awarded partial credit as this question is evaluative in nature and not application. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, real world application is not required. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness, with examples from the study to show clear understanding. These are the requirements for a Level 4 response.

PSYCHOLOGY

Paper 9990/13
Approaches, issues and debates

Key messages

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

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

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

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

Candidates need to be able to explain similarities and/or differences between studies based on psychology**.** Brief, common-sense responses can rarely be credited but these are presented by a number of candidates.

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

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

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

General comments

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

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

There were some blank responses. Candidates are encouraged to attempt all questions even if they are unsure of the answer they are providing, as they may be able to make some creditable points.

Comments on specific questions

Question 1

(a) A minority of responses correctly stated the sample size for the doodling group in the study by Andrade. Common incorrect responses included the total number of participants or the sample size for a different Core Study.

(b) Stronger responses could clearly outline the part of the procedure linked to the mock telephone message. Common points made by candidates included the word rate per minute, how long the message lasted for, and that it was played at a comfortable volume. Common incorrect responses included generic ideas about names and places in the recording. It is important for candidates to read the question carefully to ensure that they are providing the correct part of the procedure.

Question 2

(a) The majority of responses could outline the relationship stated in the question about the study by Dement and Kleitman. Many responses could identify that a positive correlation was found but then did not outline what that meant in relation to this part of the study. There were some responses that presented a response about dream duration estimation, rather than length of narrative. Some responses presented a finding from the study so could not gain any credit. It is crucial for candidates to read questions carefully.

(b) There were many clear, concise responses to this question. For example, eye movements corresponding to dream content, and estimations of length of time in REM sleep. However, there were many responses that provided a direct result from the case study. Conclusions are generic descriptions of the outcome of a study, whereas results are factual data (quantitative and/or qualitative) provided directly from the participants. Responses that focused on examples of dream content, for example, could not be credited.

Question 3

(a) There were many correct responses to this question. Popular descriptions included the sample size, that they were candidates, and an assumed high IQ. However, there was a significant minority of responses that could not describe three features/characteristics of the sample, or confused the sample with that from one of the other three groups. It is important for candidates to note the number of marks assigned to a question as this typically represents (in short answer questions) the number of correct elements that need to feature in a response. There were some blank responses to this question.

(b) A minority of responses could state the solution used by Baron-Cohen et al. to the two problems presented in the question. More candidates could note 36 pairs of eyes were used for the first problem compared to having foils of similar valence for the second problem. Common incorrect responses were using male and female eyes for problem 1 or using simpler words for problem 2.

Question 4

(a) The majority of responses could correctly identify that the children were observed in a nursery for this part of the study. A common error was to state that it happened in an experimental/observation room.

(b) The majority of responses could correctly identify that it was either an experimenter or a teacher who observed the children in this part of the study. A common error was to state it was the adult model.

(c) The minority of responses could outline how the children were rated on aggression in this part of the study. Common choices included how the rating scales were scored, and how many scales there were. Common errors included describing the final part of the study, observing children in a room with toys and the Bobo doll.

(d) There were some clear, concise responses to this question. For example, noting that the measure was a subjective assessment of aggression in the children. Other popular choices included children acting differently at school or that it was only a transient measure of aggression in the children.

There was a number of responses that did not provide an example from the study to elaborate on a chosen weakness. Candidates need to be aware that 'in this study' in the question requires an explicit example from that study to be able to access all available marks. Some responses provided a weakness of the study by Bandura et al. rather than about the baseline measure of aggression.

Question 5

(a) A minority of responses could describe a result about the Activity Index from the two groups highlighted in the question. The most popular was the placebo group scoring lower than the other group. With this type of question where the groups are already given, a description then requires a comparison and a description of what a higher/lower score indicated. Several incorrect responses chose a different result, or a result from a different part of the study. In the majority of responses, data was not provided as required in the question. It is important for candidates to read the question carefully. There were some blank responses to this question.

(b) There were some clear, concise responses to this question. For example, noting that the study had lower mundane realism, or that the sample was all male. Candidates need to be aware that 'in this study' in the question requires an explicit example from that study to be able to access all available marks. Some responses provided a weakness of the study by Schachter and Singer without any example from the study itself to tell the Examiner why it was a weakness.

Question 6

Stronger responses could clearly outline the social approach and provide clear examples from the study by Yamamoto et al. Popular examples included being influenced by the behaviour of another chimpanzee, especially if they could see them. There were a significant number of responses that were tautological and could not access marks. For example, stating that the social approach is about the social behaviour cannot be credited as it does not show sufficient understanding. To improve, candidates need to have examples from each Core Study that appropriately support each of the approaches at AS-Level.

Question 7

Reponses to this question were very varied. Stronger responses gave clear advice about the use of a model/rival technique or using rewards to shape behaviour. However, many responses focused on explaining why the advice had been given, which was not the focus of the question. The scenario set up a novel situation and told candidates about the study by Pepperberg. The question asked for an outline of advice, not to explain the reason for the advice.

Question 8

(a) The majority of responses could clearly outline what was meant by confidentiality in relation to ethical guidelines. Popular outlines included not linking data to a specific participant and not having published results as identifiable. However, some responses gave an answer about privacy, rather than confidentiality, or stated that confidentiality is 'keeping things confidential' which cannot gain credit. It is important that candidates know all ethical guidelines that psychologists must consider when designing and implementing studies.

(b) There were some strong responses provided here that fully engaged with the stimulus material. Leo was the most popular choice for the debate by a significant margin. Popular arguments supporting Leo included little protection from psychological harm and lack of informed consent. Popular arguments supporting Willow tended to focus on confidentiality. Some responses simply described aspects of ethical guidelines without engaging in the Willow-Leo debate so could only gain partial credit. To improve, candidates need to clearly choose one side of the argument and then explain why they support it using evidence from the study. Also, candidates need to know that when outlining ethical guidelines, they need to know the correct terminology (e.g., stating 'protection from psychological harm', rather than just 'protection').

Question 9

(a) A minority of responses clearly described this part of the procedure from the study by Laney et al. Common examples included rating the photographs on four scales, rating each photograph in how appetising the food looked, and that ratings tended to be on an eight-point scale. Some responses focused on a different part of the procedure or provided generic ideas about 'rating photographs'

without any specific knowledge. It is important for candidates to know the full procedures for all parts of every core study.

(b) Stronger responses could clearly explain one similarity and one difference. Popular choices to compare the studies on included ethics, the sample, laboratory studies, and experimental designs. To improve responses to this type of question, candidates need to choose comparison points that can be developed, using examples from both studies to explain the similarity and/or difference. For example, explaining the different samples can focus on the characteristics and/or sampling technique and allows for a detailed response with a statement about generalisability. However, stating that each study had a different aim does not allow the response to be detailed. There was a number of responses that compared Laney et al. to a study that was not from the cognitive approach.

Question 10

The strongest responses evaluated the study by Milgram in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of generalisability. Common choices included ecological validity, ethics, and reliability. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Milgram to explicitly support their point. These answers tended to score Level 4 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses, all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by Milgram as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. Some responses appeared to be prepared essays for Milgram, without one of the points being about generalisability. A response limited to one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did use reliability in an evaluative sense but did not fully explain why it could be a strength and/or a weakness. Several responses did not cover the named issue. In this series, more responses were attempting to focus on real world application which tended to only be awarded partial credit as this question is evaluative in nature and not application. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, real world application is not required. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples from the study to show clear understanding. These are the requirements for a Level 4 response.

PSYCHOLOGY

Paper 9990/21
Research Methods

Key messages

- This research methods paper asks candidates to answer a range of questions, including ones about the core studies in relation to research methods, terms and concepts used to describe or evaluate research methodology, and application of this knowledge to both familiar and unfamiliar contexts. Responses to this paper demonstrated a range of ability in these skills and in the application of knowledge to unfamiliar contexts.
- Candidates demonstrated excellent knowledge of the basic concept of identifying closed and open questions and good understanding of some mathematical concepts, such as bar charts and standard deviation, and methodological concepts such as structured interviews and experimental and control conditions. Some areas of basic concepts where knowledge was less evident were normal distribution, operationalisation and the description of ethical guidelines relating to animals ('species and strain' and 'numbers'). Here, candidates' responses often lacked accurate description and examples.
- The ability to link accurate detail to a given scenario or context is also required on the paper. This skill was often not always demonstrated by candidates.
- Candidates should read and follow the instructions given in the question carefully. For example, reading the whole of the stem and the question, and ensuring that they are answering the question asked.

General comments

Candidates were able to demonstrate knowledge of a range of aspects of research methods in this paper. Success was greater on simpler and low mark tariff questions such as **7(b)(ii)** and **(b)(iii)** (choosing a bar chart and using standard deviation information), **8(a)** and **8(b)** (identifying a closed and an open question – these were almost universally well done), and **9(a)** (identifying an experimental and control condition). Many candidates were also able to score some marks on more complex questions, such as other parts of **Questions 7** and **8**, but less often earned full credit. This was typically because these questions required explanation or elaboration for full marks or needed to be linked to the scenario.

Question 10(a) was rarely well answered and many responses were incomplete. Many candidates focused most of their answer on a research method other than observation, so were only able to gain marks for the small portion of their response that described observation as their data collection technique. Other candidates gave very little detail, mentioning only one aspect of their observation technique or indicating little or nothing about what behavioural responses to helping they would record. These details are essential for the description of an observational study.

Comments on specific questions

Section A

Question 1

Stronger answers referred to believing that the shocks were real or that they were really hurting someone. Many candidates repeated the essence of the stem, saying 'So they believed the situation was real'. They focused on a lack of validity if it had not looked real, without referring specifically to this therefore improving ecological validity. Other candidates correctly identified or described the avoidance of demand characteristics but offered no further detail or link to Milgram's study, so only gained limited credit.

Question 2

(a) Although many candidates were able to gain this mark, the responses were varied. There were some non-creditworthy responses including null hypotheses, further directional hypotheses and answers bringing in new variables, such as suggesting that boys would be more obedient/less obedient than girls. Many candidates were not able to demonstate a clear understanding of the meaning of 'directional' and 'non-directional'.

(b) Most candidates gave correct answers to this question part, although a few gave details of how to find the mean or median instead. A number of candidates could correctly state how to find the mode, but did not achieve the mark as they had not linked their answer to the source, e.g. 'He should find the most repeated/frequent number', without referring to 'age'. A small minority of candidates did not answer the question set, suggesting instead how age data might be collected, such as via a questionnaire.

(c) There were many generic answers to this question but where links to obedience were made, they were generally very good. Candidates tended to assume that the study would be similar to Milgram and cause stress / long term harm. Where candidates did not earn credit, this was often because they made comments regarding validity (e.g. children not understanding what the task entails) rather than comments explicitly related to ethics.

Question 3

This question did not receive many strong responses. Few candidates identified manipulated IV or measured DV, so most correct answers referred to standardisation/control. When candidates identified a feature, they often then gave an example from a different study or no example. There was repetition in answers, such as saying that laboratory experiments were done in laboratories and Schachter and Singer was an example of this as it was done in a laboratory. A number of responses also referred to other uncreditworthy ideas, such as strengths or weaknesses of laboratory experiments or to features that would be shared with other methods, such as 'has a sample'.

Question 4

(a) **Question 4** was not well answered. In this question part, very few candidates scored marks. Those who showed they understood the distribution graph (i.e. which label belongs on which axis) often did not achieve the marks as they just referred to 'percentage' on the y-axis and 'score' on the x-axis. There was also a significant number of candidates who did not attempt this question part.

(b) Similarly to **part (a)**, few candidates gained the mark on this question part, although there were more incorrect responses here than in **part (a)**.

Question 5

Many candidates gave simple and creditworthy responses to this question, with some very good suggestions for operationalisation. Weaker responses often gave a method of learning, e.g. throwing technique or modelling the skill, rather than the measurement of learning as a DV.

Question 6

Some responses to this question demonstrated very good understanding, with candidates scoring full marks. However, there were also many very low or no mark responses from candidates discussing ethics relating to humans, such as consent and right to withdraw, or animal ethics such as housing or pain and distress, rather than the two ethical issues identified in the question. There was also a lack of clarity between the use of 'species' and individual animals in many responses. A significant number of candidates appeared to misunderstand 'strain', which relates to a (genetically similar) group within species, instead raising points about distress – putting strain on animals.

Section B

Question 7

(a) (i) There were many good answers to this question, for example keeping a distance or avoiding eye contact for fear and not paying attention, crossing arms and yawning for boredom. However, many candidates needed to relate this to a questionnaire context. Candidates who did not gain credit tended to refer to how the researcher could set up a 'boring' or 'scary' situation, rather than referring to the measurement of these emotions.

(ii) Candidates produced some good answers here, for example suggesting that yawning might indicate tiredness rather than boredom, although some candidates needed to read the question more thoroughly and used an emotion in this question part which had not been used in **part (a)(i)**.

(b) (i) Most candidates could explain how to find a median, typically with a creditworthy generic explanation. Some candidates gave an incomplete answer, so could not gain credit, such as saying 'find the middle number' or 'put the numbers in a line and find the middle one'. Such responses needed to refer numerical order. A minority of incorrect responses referred to how to find a mean, a mode or a range.

(ii) Most candidates gained this mark by correctly identifying bar chart / bar graph and did not attempt to unnecessarily justify their answer.

(iii) Most candidates correctly identified empathy and many then said that this was because 3.2 was the highest, or that this emotion had the highest standard deviation. Some responses revealed some confusion in further comments, such as saying that this meant that participants showed more empathy than any other emotion.

Question 8

(a)/(b) Almost all candidates gained the marks for **Questions 8(a)** and **8(b)**.

There were very few incorrect answers, although a small number of candidates appeared to have misread the question and made up their own closed or open question. A significant number of candidates copied out the questions rather than just giving the letter as required.

(c) Most candidates referred to the problems associated with the type of data collected, e.g. quantitative data from the closed Question E or qualitative data from the open Question F. This led to many good answers, although some candidates provided only generic weaknesses which were not linked to the specific questions.

(d) (i) Most candidates correctly referred to 'set questions', 'predetermined questions', 'scripted questions', etc. A small number of candidates incorrectly stated that it was the use of closed questions which made the interview structured, so did not earn credit.

(ii) Most candidates correctly referred to the use of follow-up questions or questions based on participant responses. A small number of candidates incorrectly suggested 'adding open questions' or simply allowing participants to expand, rather than asking questions based on responses.

(e) Most candidates achieved 2 marks for correctly identifying the problem as the sample having similar characteristics and referring to this using an appropriate term, such the impact on 'generalisability' or 'validity'. Few candidates achieved full marks, as they needed elaborate to explain the effect this would have on the results.

Question 9

(a) This question part was very well answered. Only a small minority of responses were incorrect and these were typically irrelevant rather than, for example, the two conditions the wrong way round. For example, some candidates gave both the drug and the water as the experimental condition, and then some form of controlled variable as the control condition.

(b) (i) This question part was well-answered with a range of relevant ideas, such as matched containers, smell, amount, viscosity, time it was taken etc. Some candidates gave generic, non-creditworthy answers relating to sampling, such as all male, the same age etc.

(ii) There were plenty of good answers here although there were some very generic answers, e.g. 'to make it reliable', which could not be credited.

(c) Most candidates identified, or identified and explained, deception (of the control group) and potential harm (from unknown side effects or impaired therapy). Some candidates only achieved 3 marks as they only linked one of the ethical problems to the study, rather than linking each separate point to the study.

Section C

Question 10

(a) There were very few Level 3 answers. The majority of candidates designed the study as a field experiment rather than an observation, thus unnecessarily used time on groups, controls, allocation and experimental design, as this detail was irrelevant to the question. Such responses therefore lacked the information required for 'an observational study', such as covert/overt observations (the most commonly referred to) and this was often not identified explicitly and had to be inferred. Few candidates gave sufficient details of observational technique, such as including participant/non-participant observations or naturalistic/controlled observations. Some candidates did not mention what behaviours they would be recording. There were many Level 1 marks.

(b) There were some good responses to this question where aspects of the observation had been identified explicitly – for example, changing covert to overt. In contrast, when candidates had designed an experiment rather than an observation in **part 10(a)**, they gave irrelevant evaluations of experiments here so could not earn credit. Often when candidates achieved 2 marks for identifying an appropriate and specific problem, they needed to address the requirement of the question to offer a solution. When candidates did respond about their observation and attempted to tackle the issue of inter-observer reliability, they were often unsuccessful. 'Add an observer to improve inter-rater reliability' is not sufficient for credit. Inter-rater reliability is not an issue unless there are two or more observers in the first place, nor can it be improved merely by measuring reliability. The key point is to improve parity between the observers, for instance by improving operational/behavioural category definitions, by observing together and by agreeing on observations during a pilot stage. Some candidates made reference to problems with ethics or sampling which were excluded by the question.

PSYCHOLOGY

Paper 9990/22
Research Methods

Key messages

This question paper asks candidates to answer a range of questions, including ones about the core studies, in relation to research methods, terms and concepts used to describe or evaluate research methodology, and application of this knowledge to both familiar and unfamiliar contexts. These types of questions require candidates to use a variety of skills, some of which were not shown in many of the candidate responses. Candidates should be encouraged to prepare for each of these skills, especially for demonstrating knowledge of concepts and the application of this knowledge.

Ability to apply knowledge and understanding to novel scenarios is essential to help candidates to successfully complete this paper. This skill can help candidates in two ways:

Candidates should be able to apply research methods, terms and concepts to scenarios presented in questions. These can include, for example, planning, criticising or developing designs or analysing data.

Candidates should be aware of questions which require a link. When a question includes 'in this study', or makes a direct reference to the scenario, responses should go beyond simply describing or evaluating, the answer must also be contextualised in a relevant way. Practice could help candidates to learn both how to extract relevant ideas, and how to make novel suggestions based on scenarios.

Question 10 in this paper requires candidates to produce an original design for a novel research question. This 'creative' process requires practise and it is, therefore, important that candidates understand the basic research methods well and that they respond to the question by using the method stipulated by the question. Furthermore, to learn to identify flaws in a design (whether their own, as in **Question 10**, or one from a novel scenario, for example, in ***Section B***) candidates should have had the experience of practical problems in conducting studies. This is a high-level skill and can be developed through practical work with designing and conducting small studies in class, or through practice with novel scenarios. Candidates should be familiar with the overall structure of **Question 10(a)**, which can be closely tailored to requirements of an individual question, such as the required research methods and the scenario.

General comments

In general, candidate responses achieved marks across the whole range of available marks for this paper.

However, very few responses consistently and accurately demonstrated knowledge and understanding, or achieved the additional marks for linking the response to the scenarios, thus limiting marks achieved overall.

Some of the candidate responses showed a good grasp of a range of psychological concepts.

Nearly all of the questions required a link to a study. These links are of different types:
- Some links are to a specific key study.
- Other links are to the information provided in the stem (the introduction) to the question.
- A link or use of any examples from any study, as specified in the question.

If a question required a link, of any type, and that link is absent, then limited credit can be awarded.

Question 10 was rarely well answered. Candidates often began their answer by stating 'in this correlation…', but then went on to include IV, DV and features of an experiment, that were irrelevant. Candidates need to be able to plan a study for all methods that are on the syllabus and candidates should take care to read the specific requirements of the question.

Comments on specific questions

Section A

Question 1

(a) Not all candidates were able to gain credit on this question, often omitting the reason for their answer of 'directional'.

(b) In most cases any null hypothesis will begin with 'there will be no difference between' followed with the words of what is being investigated. For example: 'there will be no difference between worrying in adults with and without children'. The words 'due to chance' could be added but are not necessary. Many candidates had an answer like this and were able to gain credit. Some incorrect answers suggested that if the hypothesis is 'adults with children worry more than adults without children' then the null hypothesis must be 'adults with children do not worry more than adults without children'.

Question 2

(a) Many responses were not clear on the meaning of the term reliability; some referred to replication and others referred to validity. Credit was available for reference to consistency, for example, 'consistency of a procedure, task or measure' and a link to the study was required by the question. Many candidates were able to give an example and this was often in relation to some aspect of the procedure where the chimpanzees were always given the same tool, for example.

(b) Responses that were not able to define reliability in question **(a)** frequently struggled to achieve marks in this question part. Some candidates wrote about inter-rater reliability, however this is a measure of reliability, not how reliability can be improved as the question requested. Any improvement to procedure/task/measurement to make it more consistent such as types of chimpanzee or room conditions were awarded the available mark.

Question 3

(a) The majority of responses to this question scored 0 marks, as candidates were not always able to outline what was observed. Candidates are advised to read the words of the question carefully.

(b) (i) Most candidates were awarded the available mark for describing what is meant by qualitative data. Correct answers included describing the data as in-depth or detailed.

(ii) This question required a description of a behaviour that could have been observed to produce qualitative data. Some responses were able to identify a behaviour, such as fidgeting, flying away, eating or any verbal behaviour from Alex. Often behaviours producing quantitative data were identified, and sometimes answers were limited to 'the way Alex behaved' which needed further elaboration.

(iii) Few candidates scored credit on this question, which was linked to **(ii)**.

Question 4

(a) Many candidates provided two appropriate reasons and scored full marks. Typical answers were that being in an fMRI scanner could cause claustrophobia, whereas wearing EEG wires would not cause claustrophobia and the participant would sleep better. The question asked for a suggestion and so any appropriate answer, like the one in the above example, would receive credit. Many candidates referred to the EEG measuring eye movements, but needed to focus on brain activity as stated in the quesion.

Question 5

(a) Some candidates correctly described a median: putting scores in numerical/rank order and finding middle score. Responses then needed to apply this calculation to the Andrade study. Some candidates described how a mean or a mode could be calculated.

Question 6

This question required a description of deception and confidentiality, using examples. Answers were of three types:

- correct answers with full marks for describing both guidelines, adding detail and giving an example for each,
- answers which gave a description of each but lacked any detail or were missing examples,
- answers which were sometimes partially correct or incorrect, were missing details and examples were absent.

Question 7

(a) Candidates giving an explanation related to the study were awarded full marks, such as 'structured because she is using behaviour categories such as smiling, waving, and shouting'. Some answers incorrectly stated 'unstructured observation'. Other answers correctly stated 'structured observation' but did not provide an explanation, which was necessary for credit.

(b) (i) The majority of candidates scored full marks in response to this question. Typically, 'the participants would be unaware of her presence / would behave normally and wouldn't alter their greeting behaviour'.

(ii) This question part was answered incorrectly by most candidates who referred to a lack of consent without further elaboration. When psychologists conduct studies and influence behaviour, informed consent should be obtained from participants. Further, when observing in a private place then consent would be needed because that behaviour is private. However, when observing in a public place and when we observe what people do when they meet each other, we do not need their consent to observe them because they know they are in public and anyone could see what they do. Answers scoring full marks often referred to people moving out of view, or being too far away to clearly observe expressions.

Question 8

(a) (i)(ii) Many candidates gained full credit for these linked questions. Correct answers were for example: **(a)(i)** 'Happy – e.g. they are smiling' and for **(a)(ii)** 'Happy – for some people smiling might just be 'content' (not happy)'.

(b) Most candidates were able to gain full marks here, but there were some who did not, instead inserting numbers and drawing a bar chart, which could not be credited.

(c) (i) Any relevant ethical issue could be used, provided it was explained, as the question stated, rather than merely identified and most candidates gained full credit here.

(ii) This question required a suggestion of how the ethical problem identified in **(c)(i)** could be solved. Fewer candidates gained full credit here.

Question 9

(a) Some answers achieved full marks in response to this question, however many others needed to include more detail. The question did not have the words 'in this study' and so the answer did not have to be related to the study (Syd and sleeping).

(b) Most candidates were able to give the correct answer of an independent measures (or groups) design and the reason for this, because the epinephrine and control groups have different people in them. Some candidates incorrectly identified a repeated measures design and then were not able to explain why this was the case.

(c) A participant variable involves some feature of a participant, and a wide range of answers were provided. With most candidates scoring full marks, for example, with 'the person has drunk coffee before the study' or 'the person has insomnia', followed by consideration of how this affected Syd's experiment, such as 'caffeine might affect a person's sleep pattern', 'having insomnia would affect sleep rather than the effects of adrenaline'.

(d) A situational variable involves some feature of the situation, such as where the experiment is conducted. Many answers suggested that merely sleeping in a laboratory or sleeping in an unfamiliar bed could affect the experiment. Others suggested that the noise of being in a busy city might affect the experiment. As with **part (c)**, the answer needed to be related to Syd's experiment for the second available mark to be awarded.

Question 10

(a) The required method was a correlational study. Although candidates often began their answer by stating 'in this correlational study', details of an experiment followed with participants being allocated to conditions, and IV, DV and controls applied. Such answers were at the bottom of the mark range because candidates were not asked to design an experiment, they were not answering the question set. A correlation has only two variables. In this case, variable 1 was physical exercise and variable 2 was social media. Some candidates correctly identified these variables, but needed to measure them in a way which resulted in data that could be correlated. For example, candidates would suggest for physical exercise categorical data such as 'no exercise', '0–10 mins', '10–30 mins' and '30+ mins' which would work for a histogram but not for a scatter graph. Questions needed to be 'how many minutes do you exercise?' and the number of minutes recorded, which could then be plotted against 'how many hours do you use social media per day?' which could also be plotted. Many candidates did not measure social media, but 'screen-time' which is something very different. Screen time might be very high, but social media usage very low. For this investigation, the method used to gather the data would be in the form of a structured questionnaire, or an interview could be conducted. Some candidates suggested participants going to a laboratory, but data like this could be gathered 'in the street' or online. A large number of candidates began their answers with extensive detail about the sample, sometimes this was half the answer. Whilst details of the sample are important, they are less important than getting the details of the method correct and all the features associated with that method.

(b) There were many excellent answers. For example, candidates who gathered data using a questionnaire mentioned that the participant might give a socially desirable answer, such as saying they did much more physical exercise than they actually did, related to the plan/question. How this could be resolved was sometimes answered very well, but many candidates did not address this part of the question. Ethics and sampling were excluded by the question, but a number of answers focused on one of these aspects, which could not be credited.

PSYCHOLOGY

Paper 9990/23
Research Methods

Key messages

- This research methods paper asks candidates to answer a range of questions, including ones about the core studies in relation to research methods, terms and concepts used to describe or evaluate research methodology, and application of this knowledge to both familiar and unfamiliar contexts. Responses to this paper demonstrated a range of ability in these skills and in the application of knowledge to unfamiliar contexts.
- Candidates demonstrated excellent knowledge of basic concepts, such as finding the median and identifying closed and open questions. Many candidates would benefit from improving the skills of justifying such decisions or expanding on them to consider the implications of decisions. One area of basic concepts where such knowledge was less evident was in the description of laboratory and natural experiments. Here, candidates' responses often lacked accurate description and examples.
- The ability to link accurate detail to a given scenario or context is also required on the paper. This skill was demonstrated well by some candidates, but less well by others.
- In addition, candidates should read and follow the instructions given in the question carefully. For example, reading the whole of the stem and the question, and ensuring that they are answering the question asked.

General comments

Candidates across the ability range were able to demonstrate knowledge of a range of aspects of research methods in this paper. Success was greater on simpler and low mark tariff questions such as **1** (variables in a positive correlation), **4** (calculating a median), and **7(a)** and **(b)** (identifying a closed and an open question). These were almost universally well done. On other questions, based on novel scenarios, candidates were also successful (e.g. **Question 2(a)** and all of **Question 9**). Candidates appeared to find providing a linked answer easier in these contexts than in others so were able to apply their knowledge successfully. Most candidates were also able to score some marks on more complex questions, such as all parts of **Question 8**, but less often earned full credit. In questions such as **8(a)**, **(b)** and **(c)(i)**, explanation or elaboration was required for full marks, and in **Questions 8(c)(ii)**, **8(d)(i)** and **(ii)**, the response needed to be linked to the scenario to gain full marks.

Question 10 was sometimes very well answered although many responses were incomplete. Either the candidate used only one data collection technique or they did not mention how the results might be interpreted. These are both essential details for the description of a case study.

Comments on specific questions

Section A

Question 1

This question was well answered, with most candidates successfully identifying a positive correlation and the two co-variables. A small number of partially incorrect responses described a causal link between the two variables. Where no marks were earned, this was often because candidates were describing some aspect of the procedure or results of the study by Dement and Kleitman that were not relevant to the question.

Question 2

(a) Most candidates were able to suggest an ethical reason but not all were able to link it successfully to measuring forgetting in old people. Where candidates did not earn any marks, this was often because they had made a generic comment such as 'because of social desirability', without any indication of why this would be relevant in measuring forgetting, or to the study of very old people.

(b) There were many excellent answers to this question part. Most candidates suggested a practical reason, such as the side effects of medication or getting tired or distracted easily. However, not all were able to link this successfully to difficulties in measuring forgetting in old people. A small number of candidates gave responses that were generic and unrelated to measuring forgetting, such as 'a lack of generalisability'.

Question 3

(a) (i) Few candidates were able to answer this question part correctly. Responses were often phrased in terms of 'The item being investigated / the item used to test the aim', which could apply equally to the IV. These are the questionnaires used to collect data, so it is the measurement of the DV, rather than the manipulation on the IV, which is 'critical' in this instance.

(ii) This question part was not well answered. Many responses were simplistic, such as saying using a 'closed question' or 'rating scale'. This is 'stating' rather than 'describing' so it is not fulfilling the demands of the question. Some candidates suggested that the RQ was measured using the price a participant would pay for an item, but that is the scoring from the Food Costs Questionnaire, rather than the RQ.

(b) There were many strong answers to this question part in which candidates were able to explain the purpose of non-critical items. However, some were unable to link this to false memories or food preferences so could only earn 1 mark. Many gave elaborate detail, for example suggesting that filler questions diverted the participants from the aim, so they would be less likely to be affected by demand characteristics. This example has three generic points, but no link between this function and the context presented, i.e. the questionnaires used in the study.

Question 4

This question was very well answered. Many candidates were aware that for an accurate mean the total test scores should be divided by the number of tests the candidate sat (rather than the number of tests that had been set by the teacher).

Question 5

(a) Candidates were typically able to offer an appropriate weakness, usually poor generalisability, but this was not always well explained or linked. Few candidates went beyond 'wild chimps would be different', although those who did successfully offered both the arguments that captive chimps who had been in experiments before might be more helpful than wild ones, and that in the wild chimps may need to co-operate to survive so they would be more helpful.

(b) This question part was generally well answered, although the guideline of numbers was not always clearly identified. There were some irrelevant responses relating to matters of housing or feeding.

Question 6

There was a wide spread in answers to this question. Knowledge was stronger for laboratory experiments than for natural experiments. For the latter, descriptions often simply referred (incorrectly) to 'natural' environments' and repetitively to 'natural independent variables'. Strong responses often gave details of artificial environments based on Dement and Kleitman or Canli et al. and described the IVs in Schacter and Singer. Descriptions of controls were generally very good, although measurement of the DV was rarely given as an example. The examples offered were, however, often incorrect. Milgram was frequently cited as a laboratory experiment (it is not an experiment) and Piliavin et al. as a natural experiment (it is a field experiment). Where candidates attempted to give their own examples of natural experiments some were successful but many others were not. Those who did not earn credit for their attempts described naturalistic observations, focusing on a 'natural environment' rather than a 'naturally existing IV' and gave only one level

of an IV e.g. 'noisy trains' (with 'compared to no trains' being missing). This was often the case even when the candidate had stated that in a natural experiment the researcher cannot manipulate the IV.

Section B

Question 7

(a) (i) Almost all candidates gained the marks for **Questions 7(a)(i)** and **7(a)(ii)**. There were very few incorrect answers, although a small number of candidates appeared to misread the question and made up their own closed or open question so did not earn credit. A small number of candidates copied out the questions rather than just giving the letter as required.

(b) (i) This question part was not well answered. Where candidates gained credit, they typically referred to the problems associated with the type of data collected, e.g. quantitative data from the closed Question E, or qualitative data from the open Question F. This led to good answers, although some candidates provided only generic weaknesses which were not linked to the specific questions.

(ii) This question part was not well answered. Many candidates repeated the questions asked without offering an explanation. Where they linked the content of their answers to the aim of the study, i.e. to the type of participant (psychology candidates and the general public) responses were good, although such answers were rare. More candidates were able to earn credit by using their own knowledge of operant conditioning by explaining in a different way.

Question 8

(a) Most candidates were able to identify the correct statement as a participant variable and explain their choice, typically making comments relating to individual differences or personality. However, a significant number of candidates stated their choice without explaining it, so could not gain credit.

(b) Fewer candidates were successful in this question part than in **8(a)**. As above, many candidates needed to explain their choice. Others could not gain credit because they attempted to explain their choice by saying why it was not a participant variable, rather than why it was a situational variable. For example, comments such as 'because it did not depend on the participant's personality' were not creditworthy.

(c) (i) Most candidates were able to earn at least one mark in this question part. There were some good answers such as changing the design to matched pairs and matching each highly motivated candidate to a less motivated one. For candidates who attempted to limit either the participant or the situational variable, few were able to offer effective detail. For example, responses suggesting the participant variable could be limited by removing candidates with extreme personalities from the sample rarely suggested how this might be done (e.g. using a questionnaire to find them). Equally, other candidates suggested measuring personality but did not then go on to remove the outliers.

(ii) Those suggesting ethical problems could arise from their changes tended to present better points than those offering additional problems with validity, although some good answers of this nature were given.

(d) (i) In response to this question, many candidates offered uncreditworthy responses in terms of social desirability bias rather than demand characteristics.

(ii) A number of candidates made uncreditworthy suggestions about 'changing' the design. They suggested both implementing repeated measures (it is already, and in fact this would potentially increase the risk of demand characteristics) or independent measures (which would not work because it would worsen the problem of situational and participant variables). Stronger answers suggested various ways to distract the participants from the aim, with full credit answers identifying the technique (e.g. using filler questions on a questionnaire) and then suggesting what questions might be included.

Question 9

(a) All parts of **Question 9** were, on average, answered better than **Question 8**. Although most candidates achieved some credit here, many were simply saying 'it would not represent the population/the school' (i.e. repeating the question which does not demonstrate understanding) and relatively few were using the appropriate term, i.e. 'low generalisability'.

(b) Many candidates provided good answers to this question part. Some answers were limited to 'choose randomly', which was not sufficient for credit.

(c) (i) This question part was very well answered. However, a small number of candidates repeated 'helping', e.g. 'count how many people helped', but needed to suggest a helping behaviour.

(ii) Candidates were generally able to earn credit on this question, showing that they understood something of the nature of a non-participant observation. However, there appeared to be some confusion with covert observation, indicating that a proportion of candidates believe non-participant and covert to be synonymous. Candidates should be aware that an observer could, for example, be non-participant and overt, or participant and covert.

Section C

Question 10

(a) There were very few Level 3 answers. Many candidates used only one data collection technique, so their study was, for example, an interview or an observation, rather than a case study providing depth and breadth of detail. As a consequence, there were many Level 1 marks. Even when two (or more) techniques were used, few candidates gave sufficient details of data collection and interpretation to reach Level 3. This could have included quantitative analysis from structured observations or from repeated use of closed questions in an interview or questionnaire, and/or from the interpretation of responses to open questions. Nevertheless, some candidates did provide such information, and included excellent details about how this could be used, for example to help to improve the child's co-operative behaviour or to offer advice to the parents or teachers.

(b) In this question part, some candidates were able to provide strong answers focussing on an explicit part of their study, for example changing an interview to a questionnaire. However, even when candidates achieved two marks for identifying an appropriate and specific problem, they often needed to address the requirement of the question to offer a solution. A minority of candidates made reference to problems with ethics or sampling which were excluded by the question.

PSYCHOLOGY

Paper 9990/31
Specialist Options: Theory

Key messages

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

It is important that candidates have knowledge of the theories/explanations, terminology, and key features of studies identified in the syllabus. Some candidates were unable to identify and/or define the theories or key features of research given in these type of questions. Creating a glossary of key terms, revision of terminology/theories/studies using flash cards and class quizzes on terminology/theories/studies could prove useful. These questions are worth 2 marks and a brief response is appropriate.

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

These questions could ask the candidate to describe a theory, study or self-report used by psychologists that is named in the syllabus. These questions could also ask the candidate to describe a part of one of the named studies, such as the procedure or methods used to collect data, or a summary of the key features of the study. This question is worth 4 marks and the candidates should write a more extended answer. It would be helpful for candidates to create a revision flashcard or mind map of each bullet point in the syllabus. The flashcard should be given the title used in the syllabus, for example, Bipolar and related disorders: explanations of depression: biological: genetic and neurochemical (Oruc et al., 1997) to help the candidate identify which part of the syllabus the question is referring to as some candidates described the incorrect study. For studies, the candidate should learn the aim, sample (sampling method if known), method, procedure, two results (if possible) and conclusion.

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

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

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

This question will always come from one of the bullet points in the syllabus. Candidates could describe the three (or four) studies, theories, characteristics/explanations/treatments of disorders or techniques identified in the syllabus under the appropriate bullet point. For this exam, some of the answers used the incorrect topic area in the syllabus or the description was brief. It is possible for the responses to achieve full marks by describing at least two of the studies, theories, characteristics/explanations/treatments of disorders or techniques and this would need to be a very detailed description. It is also important that the descriptions are linked to the topic area of the syllabus. For example, **Question 8, part (a)** needed to be linked to group decision-making in organisations rather than a generic description of group decision-making. It could be useful for candidates to create revision notes with the title of each bullet point as the header. Alternatively, candidates could create a mind map and put the bullet point in the centre.

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

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

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

General comments

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

Time management for this paper was 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 abnormality were the more popular choice of option, followed by health.

Comments on specific questions

Psychology and Abnormality

Question 1

(a) There were many good responses to this question which outlined the cognitive explanation for depression. Common responses included outlining that faulty thinking can lead to depression and some gave examples of this. In addition, many described Beck's negative triad. Weaker responses frequently gave an outline of the symptoms of depression and just identified faulty thinking within the answer. Some responses gave very detailed description of the explanation which was not appropriate for a 2-mark question.

(b) There were some Level 2 responses to this question with candidates outlining the sample, something about the procedure and an indication of results. There were a few excellent responses which included the sample, procedure and details of the results. Most candidates were able to identify that Oruc et al. investigated if genes are linked to bipolar disorder. Some also identify that DNA testing was used to collect the data. A significant number of responses did not know this study and sometimes gave a description of the Gottesman and Shield's study on the genetic explanation for schizophrenia which was not creditworthy. Many responses stated that the study investigated the causes of depression rather than bipolar disorder.

(c) Candidates who knew some of the details of the Oruc et al. study in **part (b)** were able to achieve marks for this question. Common strengths included the strengths of collecting quantitative data such as being able to make comparisons/do statistical analysis and that the data was objective due to collecting DNA evidence. Weaker responses often identified the strength without giving an example to explain it. Those candidates who did not know the study sometimes achieved 1 mark by outlining a correct strength. Some responses evaluated the Gottesman and Shield's study which was not creditworthy.

Question 2

(a) Responses varied considerably for this question and covered the full range of the marks available. Some responses highlighted how well prepared some of the candidates were for this exam whereas others showed very limited knowledge of characteristics of obsessive-compulsive and related disorders including types, examples/case studies and measures. There were some responses that were detailed, accurate and coherent with a good use of psychological terminology. The best responses covered types with an outline of OCD followed by examples such as hoarding and body dysmorphic disorder. There were some very detailed descriptions of case studies such as Charles by Rappaport, and/or Lemkuhl et al.'s case study of Jason. Many responses also outlined the measures including Maudsley Obsessive-Compulsive Inventory (MOCI) and/or Yale-Brown Obsessive-Compulsive Scale (Y-BOCS). Weaker responses often gave more limited descriptions with a brief outline of the types and case studies. These responses often did not describe a measure. Some responses were confused about the difference between an obsession and a compulsion. For example, stating that someone with OCD was obsessed with hand-washing. A few responses also included reference to phobias and/or addictions which was not creditworthy.

(b) Many of the responses achieved in the Level 1 or Level 2 mark band with a few providing clear analysis and examples from **part (a)** to back up their evaluative points that enabled these type of responses to achieve Level 3 and above. There was a tendency for responses to focus on many issues per type, study and/or measure rather than evaluating issue by issue. The vast majority of responses covered the named issue of qualitative and quantitative data. However, this was frequently done in a superficial way with the response identifying a number of strengths and weaknesses of these types of data without any example from **part (a)**. This was particularly noticeable if the candidate did not outline any of the measures in **part (a)**.

Common evaluation issues included generalisability, strengths and weaknesses of case studies, application to everyday life and strengths and weaknesses of the measures. Some responses evaluated using debates such as reductionism versus holism, determinism versus free-will and individual and situational explanations. This evaluation was usually very superficial with the candidate identifying, for example, that the symptoms of the disorders given were holistic (or not) without any explanation given for why this could be considered the case.

Psychology and Consumer Behaviour

Question 3

(a) Many of the responses achieved 1 mark for giving a basic outline of an associative learning explanation for consumers' preference for product colour. Many were able to state that preferences for product colour are due to associating the product with a specific colour because of what the colour represents to the customer. Some candidates gave an example which often meant these types of responses achieved full marks. Weaker responses often just gave an example without outlining why the preference developed in the customer. A minority of responses reworded the question as their answer (i.e. 'it is the preference for a product colour due to associative learning') which was not creditworthy.

(b) There were some clear and somewhat detailed responses describing two methods used to collect data in the study by Porublev et al. study. Candidates were able to identify at least one, if not two of the methods used. Many were then able to give some details about how the method was carried out or what sort of information was collected by the researchers. Weaker responses frequently just identified the method. A common incorrect response was to identify that a questionnaire was used, rather than an interview.

(c) The marks for this question covered the full range of the mark scheme. Common strengths included detailed data, how the study met ethical guidelines and ecological validity. The most common weakness was generalisability/cultural bias. Better responses identified the strength and weakness and gave a clear example from the Porublev et al. study. Weaker responses often just identified the strength and/or weakness with either no example given or a very brief explanation. Some responses incorrectly identified that the study had quantitative data and outlined a strength and/or a weakness of this type of data which was not creditworthy.

Question 4

(a) There were a number of good and detailed descriptions of the studies by Mackay and Olshavsky, Machleit et al. and Gil et al. Weaker responses often gave fewer details of the studies with a general outline and conclusion given. Some responses gave incorrect details of the studies, such as stating that the participants in the Mackay and Olshavsky study were asked to draw maps of the supermarket rather than the route from their departure point to the supermarket. A significant minority of candidates described studies from other parts of the syllabus such as from the psychological environment: personal space which were not creditworthy.

(b) The vast majority of responses achieved Level 1 for this question. This was mainly due to the responses being very brief. Most did attempt the named issue of cultural bias and were aware that the studies were carried out in the West (although many did not know which country). Many responses focused on how overcrowding may be very common in other countries and therefore consumers would experience it differently. These responses were unable to explain how other cultures might be different to the USA in terms of shopper movement patterns or cognitive maps. Ecological validity was a common evaluation issue used but responses usually just stated that the study was carried out in 'real' supermarkets without any depth of discussion or analysis. There were a minority of good responses to this question. These responses evaluated issue by issue and gave clear examples from **part (a)** to explain their points. Some did attempt some analysis by comparing the studies in terms of the issue raised. This was frequently just identified as a similarity or a difference between the studies without explaining it.

Psychology and Health

Question 5

(a) A few of the responses were able to identify two of the measures used in the five-city project by Farquhar et al. Most common responses were to identify height, weight, blood pressure and frequency of coronary heart disease. Some candidates did not answer this question or appeared to guess at what might have been measured. Some stated how data was collected such as through a questionnaire or interview which was not creditworthy.

(b) There were many strong responses to this question with many giving a detailed description of the study on health promotion in worksites by Fox et al. Most were able to outline the sample, procedure of the study with an explanation of the token economy used as well as a brief result. Weaker responses often gave an outline of the token economy without any other details of the study. A small minority of responses did not achieve any marks due to giving an incorrect description of the study such as stating that the researchers tried to promote the health of their employees by encouraging exercise and healthy eating which was not creditworthy.

(c) Many responses identified both a strength and weakness of longitudinal research. The most common strength was that longitudinal research shows change over time. Many were able to explain this as the Fox et al. study monitored the reduction in accidents over a number of years. Common weaknesses included cost and attrition. Those responses that suggested cost as a weakness were often very brief with no example to explain why the Fox et al. study would be expensive due to it taking place over a number of years. Attrition tended to be better answered with examples given of the effects of workers leaving their job part way through the study. Some candidates gave a strength and/or weakness of the Fox et al. study (such as generalisability) which was not answering the question.

Question 6

(a) The responses to this question covered the full range of the mark scheme. Better responses gave clear and often detailed descriptions of the studies by Byrne and Long, Savage and Armstrong and Robinson and West. A few responses outlined type 1 and type 2 errors and some linked this to the effect on diagnosis and treatment. Some candidates outlined the studies on verbal communications by McKinlay and Ley and these were given credit for this question. Weaker responses were often brief and/or gave muddled descriptions of the studies which were frequently incorrect. A minority of responses gave anecdotal responses where they outlined why it is important for a practitioner to give an accurate diagnosis and ways in which they could achieve this such as being professional and make the patient feel comfortable. Some candidates described studies that were not relevant, particularly McKinstry and Wang.

(b) There were some good responses to this question. These were often able to evaluate the named issue of field experiments and were able to use the studies from **part (a)** as examples. A few responses did some good analysis by outlining the strengths and weaknesses of field experiments. Other common issues included applications to everyday life and the strengths and weaknesses of quantitative and qualitative data.

Weaker responses often covered the named issue but just outlined strengths and weaknesses of field experiments method with no examples given from **part (a)**. Some responses just identified which studies were in the field rather than evaluating the method. Many candidates chose to evaluate on a study by study basis, meaning they were unable to access higher marks as this often led to very superficial and repetitive responses.

Psychology and Organisations

Question 7

(a) There were a number of full mark responses to this question with the full equation proposed by Vroom to calculate motivation given. Weaker responses often could not label E, I and V correctly. A significant number of candidates did not know this equation and guessed that it might include rewards which was not creditworthy.

(b) Most responses were able to identify two non-monetary rewards that are motivators at work. Common answers included recognition and praise and these were frequently well described. Weaker responses sometimes just identified the rewards without any details/example of the reward or gave a monetary reward such as vacations, bonuses, or company cars which was not creditworthy.

(c) Most responses were able to achieve Level 1 or Level 2 for this question. The most common strength was that non-monetary rewards save the organisation money and also lead to a motivated workforce which can increase profits. Common weaknesses included individual differences, resentment from other workers seeing a colleague receiving a reward and praise/respect do not pay the employee's bills. Weaker responses were brief and did not explain the strength and/or weakness. Some responses evaluated the monetary reward they had put in **part (a)** which was not creditworthy.

Question 8

(a) There were a few good, detailed responses to this question. Some candidates displayed good knowledge of Wedley and Field's theory about the decision-making process, groupthink and how to avoid it, as well as an outline of cognitive limitations and errors. The best description was of groupthink and strategies to prevent it when groups are formed to make decisions in organisations. A significant number of responses were anecdotal: for example, some outlined why making good decisions is important in an organisation. Some candidates were able to outline groupthink but suggested it is positive for decision-making which is incorrect.

(b) There were a few good responses to this question. These candidates often started their response with the named issue of practical applications and were able to explain how these theories about decision-making can be used by organisations to improve decisions as well as strategies to avoid groupthink. Most candidates did not do analysis for this issue where they could have outlined why it would be difficult for an organisation to implement these theories in practice. Other common

evaluation issues included cultural bias and individual and situational explanations. Weaker responses tended to be very superficial with many just redescribing the theories outlined in **part (a)** for practical applications. Some candidates evaluated the theories as though they were studies and used issues such as ecological validity and reliability which were not creditworthy.

PSYCHOLOGY

Paper 9990/32
Specialist Options: Theory

Key messages

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

It is important that candidates are made aware of the terminology, theories, disorders and techniques identified in the syllabus as some were unable to identify and/or define the terms/concepts given in these types of questions. Creating a glossary of key terms, revision of terminology/theories using flash cards and class quizzes on terminology/theories could prove useful. These questions are worth 2 marks and a brief response is appropriate.

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

These questions could ask the candidate to describe a theory, study or self-report used by psychologists that is named in the syllabus. These questions could also ask the candidate to describe a part of one of the named studies, such as the procedure or results, or a summary of the key features of the study. This question is worth 4 marks and the candidates should write a more extended answer. It would be helpful for candidates to create a revision flashcard or mind map of each bullet point in the syllabus. The flashcard should be given the title used in the syllabus, for example, Adherence to medical advice: measuring non-adherence: objective: pill counting (Chung and Naya, 2000) to help the candidate identify which part of the syllabus the question is referring to as some candidates described the incorrect study or theory. For studies, the candidate should learn the aim, sample (sampling method if known), method, procedure, two results (if possible) and conclusion.

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

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

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

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

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

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

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

General comments

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

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

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

Comments on specific questions

Psychology and Abnormality

Question 1

(a) There were many good responses to this question which outlined the symptoms of kleptomania and/or identified it as an impulse control disorder. Common responses included outlining that it is the urge to steal and objects stolen often are not needed/have little monetary value. 1 mark responses frequently just stated that kleptomania is the urge to steal. Some responses stated that patients with kleptomania experience pleasure when they steal which was not creditworthy as a thief would also experience pleasure when stealing. Better responses instead stated that patients with kleptomania experience tension prior to stealing and feel relief afterwards. Some responses incorrectly outlined pyromania.

(b) The majority of responses were able to outline at least one of the components that Griffiths used to define addiction and many outlined two. Common responses included salience, mood modification, withdrawal and conflict. Good responses both identified and outlined the two components. Weaker responses frequently just identified the component without a description or gave an incorrect outline. Withdrawal and tolerance were sometimes incorrectly outlined by candidates. Some responses outlined what is meant by substance and non-substance addiction which was not creditworthy.

(c) The marks for this question covered the full range of the mark scheme. Common strengths included that the components provide a comprehensive definition of addiction as well as practical applications for diagnosis and treatment. Some candidates linked these two points together which helped them to achieve high marks. Common weaknesses included individual differences (not every addict will experience all of the components), reductionism as biological/genetic factors are not considered and cultural differences. Good responses outlined the strength/weakness and used one or more of the components to explain the point. Weaker responses were often brief or did not give any example from the components. Some responses confused the components with a study and evaluated the ecological validity of it which was not creditworthy.

Question 2

(a) Responses varied for this question and covered the full range of the marks available. Some responses highlighted how well prepared some of the candidates were for this exam whereas others showed very limited knowledge of treatments for schizophrenia. The best responses outlined three of the treatments and included details of the studies by Paul and Lentz and Sensky. Token economies / Paul and Lentz's study were described in the most depth. Cognitive behaviour therapy was the weakest description in most responses with many just identifying that it is a talking therapy and challenges the irrational thoughts of the patients, without linking the description to any of the symptoms of schizophrenia. Weaker responses often gave brief descriptions and inaccurate details of some of the treatments. Inaccuracies were often seen with electro-convulsive therapy and antipsychotics. In addition, many candidates gave the side effects of treatments which is not creditworthy in **part (a)** of this question and should be given in **part (b)**. Some candidates gave lengthy descriptions of the symptoms of schizophrenia which was not creditworthy. In addition, some responses described the study by Lovell et al. which is a cognitive therapy for obsessive-compulsive disorder and is also not creditworthy for this question.

(b) Many of the responses achieved in the Level 1 or Level 2 mark band with a few providing clear analysis and examples from **part (a)** to back up their evaluative points that enabled these types of responses to achieve Level 3 and above. The vast majority of responses covered the named issue of ethics. There were some well argued responses with clear comparisons made between treatments. Weaker responses argued that antipsychotics and ECT are unethical due to side effects without considering that most patients consent to the treatment and it does improve their symptoms and quality of life. Other issues discussed included appropriateness of the treatments, determinism versus free-will and nature versus nurture. There was a tendency for responses to focus on many issues per treatment rather than considering each issue in turn which would have enabled the candidate to provide analysis.

Psychology and Consumer Behaviour

Question 3

(a) There were some good responses to this question. Full mark responses outlined that 'satisficing' is where the consumer decides that the product meets their basic requirements. The consumer then stops their decision-making and purchases the product. 1 mark responses usually outlined that the consumer will want a product that meets their needs but these responses did not clearly link to decision-making. Many responses stated that consumers choose the best product or that the consumer was 'satisfied' with the product which were not creditworthy answers to this question.

(b) There were a few good responses to this question. Common features of prospect theory given described how consumers value gains and losses differently with an example. Some responses identified value and endowment and a few were able to explain what these terms mean. Many candidates did not know prospect theory and often outlined the effect of the store environment on purchases or how companies can encourage consumers to purchase their products. These types of responses were not creditworthy.

(c) Candidates who achieved some marks in **part (b)** for their description of prospect theory were able to achieve marks for this question. Common weaknesses included individual differences and reductionism as factors other than value and endowment will have an impact on purchasing decisions. Most creditworthy responses did tend to be brief and often achieved Level 1 or Level 2. Some responses outlined weaknesses such as ecological validity and generalisability and incorrectly referred to prospect theory as a study. These types of responses were not creditworthy.

Question 4

(a) There were some good responses to this question. Some candidates provided impressive details of the studies by Kardes et al. as well as good details of either Cialdini's six ways to close a sale or Cardone's steps to close a sale. Some responses did give clear details of customer-/competitor-/ product-focused sales techniques. Weaker responses often gave fewer details of the Kardes et al. study and just a few ways to close a sale. It was common for the sales techniques to not be included in weaker responses. A significant number of responses described advertising, discounts or gave an anecdotal response outlining what a company could do to increase sales which was not creditworthy.

(b) A significant number of candidates did not answer this question. Those that did attempt it often achieved very low marks as they had not been able to describe the techniques for selling the product in **part (a)**. Many responses continued to outline what a company could do to increase sales as an attempt at the named issue of usefulness. The candidates who achieved marks in **part (a)** were able to access some marks for this question. Most attempted the named issue of usefulness but frequently just stated that the findings of the Kardes et al. study and the selling techniques were useful without considering why or how they might not be useful to a company. For example, how does a company decide whether to use customer-, competitor- or product-focused sales techniques or why do the steps to close a sale work effectively for some customers and not for others. Other common issues included evaluating the Kardes et al. study using issues such as generalisability, ecological validity, strengths and weaknesses of the data collection methods and ethics.

Psychology and Health

Question 5

(a) Many of the candidates achieved full marks for this question by identifying blood and urine tests. Some responses were lengthy for a two mark question that asked the candidates to identify two tests. A common, incorrect response was to identify pill counting.

(b) There were many strong responses to this question with many giving a detailed description of the procedure of the study on pill counting by Chung and Naya. Common details given were the sample, the instructions for taking the medication, details of the TrackCap and (less commonly) what was done at the end of the study as a check. Weaker responses gave fewer points regarding the procedure with some just outlining how the TrackCap worked. Some responses only stated that the researchers counted the pills of the participants.

(c) There were a number of good responses to this question and the vast majority were able to give at least one weakness of the Chung and Naya study. The most common weakness given was the lack of validity as the TrackCap records when the bottle is open and not if the pill was taken. There were some good responses as well for a lack of generalisability of the sample (citing number of participants or that all patients suffered from asthma). Weaker responses frequently repeated the weakness of the accuracy of the TrackCap which was credited once. In addition, some responses used incorrect terminology such as stating that the study lacked ecological validity due to the problem of not knowing if the patient took their pill after opening the TrackCap. These types of responses often achieved Level 1 or Level 2. Responses that were not awarded marks were frequently those that just identified an issue or debate such as reductionism without any explanation.

Question 6

(a) The responses to this question covered the full range of the mark scheme. Better responses gave clear and often detailed descriptions of the study by Chandola et al. and the Holmes and Rahe life events scale. Some of the responses also outlined Rahe's study on US sailors which showed the link between life events and ill health. Many responses did describe the GAS model but frequently just outlined the 'fight or flight' response without linking their description to stress and/or the three stages of alarm, resistance and exhaustion. There was a variety of quality of response in the description of the effect of personality on stress and ill health. Some responses gave clear descriptions of Type A and B personalities and explained why Type A would experience more stress as well as sometimes outlining the longitudinal study done by Friedman and Rosenman to monitor the health outcomes of both personality types. Weaker responses frequently just identified

the features of Type A and B personalities without any link to stress or health. There was a significant number of responses that were generic. There were some very long responses that just included irrelevant, anecdotal description of stress as a concept and how it might be related to life, work, exams and tenuous links to the effects on health. These types of responses were not creditworthy.

(b) Most responses achieved Level 1 or Level 2 for this question. Most did the named issue of nature versus nurture and a few good responses were seen that were able to explain clearly why the causes of stress outlined in **part (a)** supported one side of the debate. A few responses did analysis where they were able to explain how one of the causes could be seen to support both sides of the debate. For example, the GAS model supports nature as the 'fight or flight' / alarm stage is automatic and the body's natural physiological reaction to dangerous events. However, it does also support nurture as people have learned to have an alarm response to stressful events which are not physically dangerous. Other common issues included practical applications, determinism versus free-will, situational and individual explanations and strengths and weaknesses of the studies outlined in **part (a)** with evaluation of the Chandola et al. study being the most common. Weaker responses tended to provide a superficial discussion where the candidate merely identified which side of a debate each cause supported without any explanation given for this.

Psychology and Organisations

Question 7

(a) Many responses were able to achieve 1 mark for this question. Common responses included outlining that organisational commitment involves believing in the goals/values of the company or loyalty to the organisation. Some responses also included reference to one or more of Allen and Meyer's three types of organisational commitment (continuance, affective, normative). Weaker responses often just mentioned loyalty. Responses that just stated that commitment involves being committed to the organisation could not be credited.

(b) The vast majority of responses were able to identify at least one and often two hygiene factors identified by Herzberg in the two-factor theory of job satisfaction. The most common factors were salary and working conditions. Many responses gave an outline of the factors identified and/or gave an example. Weaker responses frequently just identified the hygiene factors rather than outlining them. Some responses did not describe the factor but instead stated that it would affect job satisfaction. There was confusion for some candidates about the working conditions where some stated that it was whether the work environment was comfortable, rather than being safe. A few responses appeared to take the word hygiene literally and talked about washing themselves which was not creditworthy. A small number of responses incorrectly outlined motivational factors such as recognition and responsibility.

(c) Most responses were able to achieve Level 1 or Level 2 for this question. Common strengths included the usefulness of the theory with examples of how organisations could use the theory to improve satisfaction and holistic/comprehensive theory. Common weaknesses included individual differences and weaknesses of the study by Hertzberg in Pittsburgh. Better responses identified the strength/weakness and gave an example from Hertzberg's theory to explain the point. An example of this is where the candidate has identified usefulness and then gives two examples of how an organisation could ensure that they provide for the motivational and hygiene factors to improve satisfaction. Weaker responses often lacked detail. Some responses stated that the theory was not backed up by any research which did not receive any credit.

Question 8

(a) There were many good, detailed responses to this question. Maslow's hierarchy of needs was often described in depth with many outlining both the five and eight tier hierarchy. Good responses gave examples of each of the stages and explained that lower level needs must be satisfied first before someone moves to the next tier. The descriptions of ERG theory tended to be more succinct and many responses outlined that this theory was not hierarchical. A minority of responses confused what the three letters in the theory stand for. There were a variety of details given about McClelland's achievement motivation theory with some responses giving a lot of detail and clearly explaining the three work-related needs. Weaker responses often just identified the tiers/categories in the theories without any descriptions given of them. In addition, weaker responses often did not

explain the hierarchical nature of Maslow's theory (or compare it to the non-hierarchical theory by Alderfer).

(b) The marks for this question were commonly in Level 1 and Level 2. Most responses attempted the named issue of determinism versus free-will and some were able to give a clear explanation as to which side of this debate each of the need theories supports. Many of the responses were very superficial for this issue and merely stated that each theory supported determinism (or free-will) with no explanation given. Other common evaluation issues included application to everyday life, reductionism versus holism and individual and situational explanations. Most of the responses evaluated each need theory in turn rather than on an issue-by-issue basis so discussion was rather superficial. These types of responses were awarded a Level 1 mark.

PSYCHOLOGY

Paper 9990/33
Specialist Options: Theory

Key messages

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

It is important that candidates have knowledge of the theories/explanations, terminology, and key features of studies identified in the syllabus. Some candidates were unable to identify and/or define the theories or key features of research given in these type of questions. Creating a glossary of key terms, revision of terminology/theories/studies using flash cards and class quizzes on terminology/theories/studies could prove useful. These questions are worth 2 marks and a brief response is appropriate.

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

These questions could ask the candidate to describe a theory, study or self-report used by psychologists that is named in the syllabus. These questions could also ask the candidate to describe a part of one of the named studies, such as the procedure or methods used to collect data, or a summary of the key features of the study. This question is worth 4 marks and the candidates should write a more extended answer. It would be helpful for candidates to create a revision flashcard or mind map of each bullet point in the syllabus. The flashcard should be given the title used in the syllabus, for example, Bipolar and related disorders: explanations of depression: biological: genetic and neurochemical (Oruc et al., 1997) to help the candidate identify which part of the syllabus the question is referring to as some candidates described the incorrect study. For studies, the candidate should learn the aim, sample (sampling method if known), method, procedure, two results (if possible) and conclusion.

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

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

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

This question will always come from one of the bullet points in the syllabus. Candidates could describe the three (or four) studies, theories, characteristics/explanations/treatments of disorders or techniques identified in the syllabus under the appropriate bullet point. For this exam, some of the answers used the incorrect topic area in the syllabus or the description was brief. It is possible for the responses to achieve full marks by describing at least two of the studies, theories, characteristics/explanations/treatments of disorders or techniques and this would need to be a very detailed description. It is also important that the descriptions are linked to the topic area of the syllabus. For example, **Question 8, part (a)** needed to be linked to group decision-making in organisations rather than a generic description of group decision-making. It could be useful for candidates to create revision notes with the title of each bullet point as the header. Alternatively, candidates could create a mind map and put the bullet point in the centre.

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

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

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

General comments

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

Time management for this paper was 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 abnormality were the more popular choice of option, followed by health.

Comments on specific questions

Psychology and Abnormality

Question 1

(a) There were many good responses to this question which outlined the cognitive explanation for depression. Common responses included outlining that faulty thinking can lead to depression and some gave examples of this. In addition, many described Beck's negative triad. Weaker responses frequently gave an outline of the symptoms of depression and just identified faulty thinking within the answer. Some responses gave very detailed description of the explanation which was not appropriate for a 2-mark question.

(b) There were some Level 2 responses to this question with candidates outlining the sample, something about the procedure and an indication of results. There were a few excellent responses which included the sample, procedure and details of the results. Most candidates were able to identify that Oruc et al. investigated if genes are linked to bipolar disorder. Some also identify that DNA testing was used to collect the data. A significant number of responses did not know this study and sometimes gave a description of the Gottesman and Shield's study on the genetic explanation for schizophrenia which was not creditworthy. Many responses stated that the study investigated the causes of depression rather than bipolar disorder.

(c) Candidates who knew some of the details of the Oruc et al. study in **part (b)** were able to achieve marks for this question. Common strengths included the strengths of collecting quantitative data such as being able to make comparisons/do statistical analysis and that the data was objective due to collecting DNA evidence. Weaker responses often identified the strength without giving an example to explain it. Those candidates who did not know the study sometimes achieved 1 mark by outlining a correct strength. Some responses evaluated the Gottesman and Shield's study which was not creditworthy.

Question 2

(a) Responses varied considerably for this question and covered the full range of the marks available. Some responses highlighted how well prepared some of the candidates were for this exam whereas others showed very limited knowledge of characteristics of obsessive-compulsive and related disorders including types, examples/case studies and measures. There were some responses that were detailed, accurate and coherent with a good use of psychological terminology. The best responses covered types with an outline of OCD followed by examples such as hoarding and body dysmorphic disorder. There were some very detailed descriptions of case studies such as Charles by Rappaport, and/or Lemkuhl et al.'s case study of Jason. Many responses also outlined the measures including Maudsley Obsessive-Compulsive Inventory (MOCI) and/or Yale-Brown Obsessive-Compulsive Scale (Y-BOCS). Weaker responses often gave more limited descriptions with a brief outline of the types and case studies. These responses often did not describe a measure. Some responses were confused about the difference between an obsession and a compulsion. For example, stating that someone with OCD was obsessed with hand-washing. A few responses also included reference to phobias and/or addictions which was not creditworthy.

(b) Many of the responses achieved in the Level 1 or Level 2 mark band with a few providing clear analysis and examples from **part (a)** to back up their evaluative points that enabled these type of responses to achieve Level 3 and above. There was a tendency for responses to focus on many issues per type, study and/or measure rather than evaluating issue by issue. The vast majority of responses covered the named issue of qualitative and quantitative data. However, this was frequently done in a superficial way with the response identifying a number of strengths and weaknesses of these types of data without any example from **part (a)**. This was particularly noticeable if the candidate did not outline any of the measures in **part (a)**.

Common evaluation issues included generalisability, strengths and weaknesses of case studies, application to everyday life and strengths and weaknesses of the measures. Some responses evaluated using debates such as reductionism versus holism, determinism versus free-will and individual and situational explanations. This evaluation was usually very superficial with the candidate identifying, for example, that the symptoms of the disorders given were holistic (or not) without any explanation given for why this could be considered the case.

Psychology and Consumer Behaviour

Question 3

(a) Many of the responses achieved 1 mark for giving a basic outline of an associative learning explanation for consumers' preference for product colour. Many were able to state that preferences for product colour are due to associating the product with a specific colour because of what the colour represents to the customer. Some candidates gave an example which often meant these types of responses achieved full marks. Weaker responses often just gave an example without outlining why the preference developed in the customer. A minority of responses reworded the question as their answer (i.e. 'it is the preference for a product colour due to associative learning') which was not creditworthy.

(b) There were some clear and somewhat detailed responses describing two methods used to collect data in the study by Porublev et al. study. Candidates were able to identify at least one, if not two of the methods used. Many were then able to give some details about how the method was carried out or what sort of information was collected by the researchers. Weaker responses frequently just identified the method. A common incorrect response was to identify that a questionnaire was used, rather than an interview.

(c) The marks for this question covered the full range of the mark scheme. Common strengths included detailed data, how the study met ethical guidelines and ecological validity. The most common weakness was generalisability/cultural bias. Better responses identified the strength and weakness and gave a clear example from the Porublev et al. study. Weaker responses often just identified the strength and/or weakness with either no example given or a very brief explanation. Some responses incorrectly identified that the study had quantitative data and outlined a strength and/or a weakness of this type of data which was not creditworthy.

Question 4

(a) There were a number of good and detailed descriptions of the studies by Mackay and Olshavsky, Machleit et al. and Gil et al. Weaker responses often gave fewer details of the studies with a general outline and conclusion given. Some responses gave incorrect details of the studies, such as stating that the participants in the Mackay and Olshavsky study were asked to draw maps of the supermarket rather than the route from their departure point to the supermarket. A significant minority of candidates described studies from other parts of the syllabus such as from the psychological environment: personal space which were not creditworthy.

(b) The vast majority of responses achieved Level 1 for this question. This was mainly due to the responses being very brief. Most did attempt the named issue of cultural bias and were aware that the studies were carried out in the West (although many did not know which country). Many responses focused on how overcrowding may be very common in other countries and therefore consumers would experience it differently. These responses were unable to explain how other cultures might be different to the USA in terms of shopper movement patterns or cognitive maps. Ecological validity was a common evaluation issue used but responses usually just stated that the study was carried out in 'real' supermarkets without any depth of discussion or analysis. There were a minority of good responses to this question. These responses evaluated issue by issue and gave clear examples from **part (a)** to explain their points. Some did attempt some analysis by comparing the studies in terms of the issue raised. This was frequently just identified as a similarity or a difference between the studies without explaining it.

Psychology and Health

Question 5

(a) A few of the responses were able to identify two of the measures used in the five-city project by Farquhar et al. Most common responses were to identify height, weight, blood pressure and frequency of coronary heart disease. Some candidates did not answer this question or appeared to guess at what might have been measured. Some stated how data was collected such as through a questionnaire or interview which was not creditworthy.

(b) There were many strong responses to this question with many giving a detailed description of the study on health promotion in worksites by Fox et al. Most were able to outline the sample, procedure of the study with an explanation of the token economy used as well as a brief result. Weaker responses often gave an outline of the token economy without any other details of the study. A small minority of responses did not achieve any marks due to giving an incorrect description of the study such as stating that the researchers tried to promote the health of their employees by encouraging exercise and healthy eating which was not creditworthy.

(c) Many responses identified both a strength and weakness of longitudinal research. The most common strength was that longitudinal research shows change over time. Many were able to explain this as the Fox et al. study monitored the reduction in accidents over a number of years. Common weaknesses included cost and attrition. Those responses that suggested cost as a weakness were often very brief with no example to explain why the Fox et al. study would be expensive due to it taking place over a number of years. Attrition tended to be better answered with examples given of the effects of workers leaving their job part way through the study. Some candidates gave a strength and/or weakness of the Fox et al. study (such as generalisability) which was not answering the question.

Question 6

(a) The responses to this question covered the full range of the mark scheme. Better responses gave clear and often detailed descriptions of the studies by Byrne and Long, Savage and Armstrong and Robinson and West. A few responses outlined type 1 and type 2 errors and some linked this to the effect on diagnosis and treatment. Some candidates outlined the studies on verbal communications by McKinlay and Ley and these were given credit for this question. Weaker responses were often brief and/or gave muddled descriptions of the studies which were frequently incorrect. A minority of responses gave anecdotal responses where they outlined why it is important for a practitioner to give an accurate diagnosis and ways in which they could achieve this such as being professional and make the patient feel comfortable. Some candidates described studies that were not relevant, particularly McKinstry and Wang.

(b) There were some good responses to this question. These were often able to evaluate the named issue of field experiments and were able to use the studies from **part (a)** as examples. A few responses did some good analysis by outlining the strengths and weaknesses of field experiments. Other common issues included applications to everyday life and the strengths and weaknesses of quantitative and qualitative data.

Weaker responses often covered the named issue but just outlined strengths and weaknesses of field experiments method with no examples given from **part (a)**. Some responses just identified which studies were in the field rather than evaluating the method. Many candidates chose to evaluate on a study by study basis, meaning they were unable to access higher marks as this often led to very superficial and repetitive responses.

Psychology and Organisations

Question 7

(a) There were a number of full mark responses to this question with the full equation proposed by Vroom to calculate motivation given. Weaker responses often could not label E, I and V correctly. A significant number of candidates did not know this equation and guessed that it might include rewards which was not creditworthy.

(b) Most responses were able to identify two non-monetary rewards that are motivators at work. Common answers included recognition and praise and these were frequently well described. Weaker responses sometimes just identified the rewards without any details/example of the reward or gave a monetary reward such as vacations, bonuses, or company cars which was not creditworthy.

(c) Most responses were able to achieve Level 1 or Level 2 for this question. The most common strength was that non-monetary rewards save the organisation money and also lead to a motivated workforce which can increase profits. Common weaknesses included individual differences, resentment from other workers seeing a colleague receiving a reward and praise/respect do not pay the employee's bills. Weaker responses were brief and did not explain the strength and/or weakness. Some responses evaluated the monetary reward they had put in **part (a)** which was not creditworthy.

Question 8

(a) There were a few good, detailed responses to this question. Some candidates displayed good knowledge of Wedley and Field's theory about the decision-making process, groupthink and how to avoid it, as well as an outline of cognitive limitations and errors. The best description was of groupthink and strategies to prevent it when groups are formed to make decisions in organisations. A significant number of responses were anecdotal: for example, some outlined why making good decisions is important in an organisation. Some candidates were able to outline groupthink but suggested it is positive for decision-making which is incorrect.

(b) There were a few good responses to this question. These candidates often started their response with the named issue of practical applications and were able to explain how these theories about decision-making can be used by organisations to improve decisions as well as strategies to avoid groupthink. Most candidates did not do analysis for this issue where they could have outlined why it would be difficult for an organisation to implement these theories in practice. Other common

evaluation issues included cultural bias and individual and situational explanations. Weaker responses tended to be very superficial with many just redescribing the theories outlined in **part (a)** for practical applications. Some candidates evaluated the theories as though they were studies and used issues such as ecological validity and reliability which were not creditworthy.

PSYCHOLOGY

Paper 9990/41
Specialist Options: Application

Key messages

(a) What has been learned from the AS component of the syllabus should be transferred to the A2 component. For example, at AS candidates learn about methodology, such as experiments, which also apply to A2.

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

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

(d) In ***Section B***, **Questions 5**, **6**, **7** and **8**, methodological knowledge must be evident and detailed for top marks to be accessed. The procedure, however detailed is just one methodological aspect. For top marks answers must explain methodology rather than merely identify it.

(e) In ***Section C***, **Questions 9**, **10**, **11** and **12**, to access top marks, answers must include a debate which has two sides, such as strengths/advantages and weaknesses/disadvantages. Supporting evidence should also be provided. Description cannot be credited.

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

General comments

Some candidates appeared not to have studied two options at the same level of depth, and a few candidates answered questions from one option only. Some candidates answered questions from three and even four options. Whilst answers to one option were often very good, some answers to the second option were very poor, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

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

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

(iii) For **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so marks were limited.
Candidates should not use terms without explanation. Frequently answers stated 'it is reductionist' or 'it is useful in everyday life' without further explanation. To state 'it is reductionist' is merely to identify; it is not automatically a strength or weakness. Further many candidates assume that to be reductionist is always a weakness. It is not; any experiment is reductionist because variables are controlled and only the IV is manipulated. Reductionism is the basis of any experiment and as such it is a strength.

(iv) Many conclusions repeated what had already been written, and such summaries scored no marks. A conclusion is a 'decision reached by reasoning' and so as the reasoning has been done through the advantages and disadvantages, a final decision/conclusion needs to be drawn.

(v) Candidates should think about what the question requires rather than writing pre-prepared answers. Many questions will test the ability to apply knowledge from one thing to another, particularly methodological knowledge.

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

Section B

Many candidates conduct an experiment whatever the question. An interview, questionnaire or observation are methods independent of an experiment and candidates should not try to make other methods 'fit' into an experimental format. Answers to **part (a)** questions in this section should include an appropriate design, have applied a range (four or five) of relevant methodological design features, each of which should be explained fully, showing good understanding. Many answers listed features such as 'I would have a random sample' and 'It would be an independent measures design' without explanation of why it would be a random sample, or how this would be obtained.

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

Section C

It is essential that answers focus on the question that is set. Every question in this Section invites candidates to consider the extent to which they agree or disagree with the statement, rather than describe what they know about that topic area, and answers that do not address the question will only achieve minimal marks. To score marks at the top end of the mark range, answers must focus on arguments both for and against the statement, answers must the use appropriate evidence to support the argument, and, at the very top of the mark range, answers should show awareness of wider issues and evidence that is relevant.

Comments on specific questions

Section A

Question 1

(a) This question asked for a behavioural explanation of phobias, and nearly all answers were awarded 1 mark for stating that 'all behaviours are learned'. Some answers went further and explained how this was done and some answers, also awarded 2 marks, gave the example of little Albert who was conditioned to be afraid of a white rat.

(b) There were some excellent answers in response to this question, but there were also some zero mark answers. The strongest answers included the two essential elements: the creation of an anxiety hierarchy and the learning of progressive muscle relaxation. Inclusion of these two elements scored 2 marks and for the 'in this study' component, relating Richard's fear of oranges to systematic desensitisation, scored the additional 2 marks.

(c) Two other ways to treat a phobia that are included on the syllabus are Ost's applied tension (to treat blood/injection phobia) and cognitive behaviour therapy (Ost and Westling) which can be used to treat any phobia. Many candidates opted for these two, but other possibilities were also acceptable in this 'suggest' question. For example, psychoanalytic psychotherapy was sometimes mentioned.

(d) When considering the strengths and weaknesses of using systematic desensitisation most candidates provided logical and appropriate responses. However, a few answers needed to acknowledge that people undergoing systematic desensitisation are not participants in a study but real people with real phobias. Comments such as 'participants might try to move through the hierarchy quicker' are incorrect; the opposite is often the case. Similarly, 'the participant might be harmed by the treatment, so it is unethical' is also incorrect because the person has chosen the treatment to reduce the effects of the phobia and also because the hierarchy is designed so it is not distressing or harmful.

Question 2

(a) There were three types of answer in response to this question: those who had no understanding of the retroactive effect (or interference); those who confused it with proactive interference; and those who knew and understood the term. Those who related retroactive interference to the Braun-LaTour study were awarded 2 marks.

(b) (i) Many candidates were not able to give an example of the data collected, and are advised to carefully read the stem of the question which prompts the response.

(ii) Most candidates were awarded full marks. In this instance there was no requirement to relate it to the study, although this was creditable as explanation.

(c) The situational component would involve the past experience of visiting Disneyland and whether they had seen certain characters there, such as Bugs Bunny. The individual component involved the false memory (a cognitive process) that was created in the mind of the participant. Some excellent answers focused clearly on these two and showed good understanding.

(d) Many answers included two strengths and two weaknesses of giving course credits to students, but often answers only scored partial marks because answers were not related to the study of false advertising as the question required. A simple formula applies to all **part (d)** questions: strength + example, strength + example, weakness + example, weakness + example. Note that **(i)** 1 mark required strength or weakness *plus* an example; **(ii)** the question required strengths and weaknesses (plural) not just one of each.

Question 3

(a) Most candidates could not outline an appropriate conclusion from the Riekert and Drotar study. Two conclusions appear in the abstract of the Riekert and Drotar study and either could be credited.

(b) (i) Many suggested that the children were under 16 years and could not give consent. Notably in any study an adult may give consent for a child when the child does not wish it. In this study the researchers required consent from *both* parent and child. Correct answers were therefore that 14 of the families (both adolescent and parent) did not give consent to participate. Another correct answer reducing the number of participants was that 28 participants (labelled 'non-returners') did not return the postal questionnaire as they applied their right to withdraw.

(ii) There was a wide range of incorrect answers to this question. Excluding participants by researchers from any study is important in psychological research because it might be that participants might be harmed by the research, for example. Given the nature of the study, Riekert and Drotar excluded participants because: they had been diagnosed with insulin-dependent diabetes mellitus or because they had an additional chronic illness or because they showed a developmental delay such as Down's Syndrome.

(c) Most answers were awarded maximum marks for explaining two reasons why a patient might not adhere to medical advice. Rational non-adherence was given by most candidates and factors from the health belief model were also appropriate. A few candidates provided 'the cost of the medicine'

which was awarded 1 mark because it is not incorrect but needed more detail for it to be awarded another mark.

(d) There were many answers related to the strengths and weaknesses of questionnaires in general, such as 'people might not give honest answers' but this question was about 'using questionnaires to be returned by post' which is rather more specific. Candidates should always read the question carefully and address the focus in their answer. Some answers did address the specific question but needed to relate their answers to gathering data about health as the question required.

Question 4

(a) Most candidates correctly outlined a 'rotator', writing 'a worker who works a pattern of rotating (alternating) shifts of days/evenings/nights' and adding, for example 'compared to day workers, evening workers, night workers and day workers with occasional nights' for the extra detail.

(b) Nearly all candidates scored 1 mark for one finding, but the question required two findings. Some answers were correct but needed to have more detail. Writing 'night workers had less than 5 hours sleep' is correct and scored 1 mark. Writing '20% of night workers had less than 5 hours sleep compared to other groups such as rotators' scored 2 marks.

(c) (i) This question required candidates to identify two types of error or accident more likely to happen to nurses on a rotator shift. Some candidates did not address the latter part of the question and gave two generalised responses. Others successfully identified two of the four possibilities.

(ii) This question required candidates to suggest two effects shiftwork can have on health, and they could have used the findings of the Knuttson (2003) study, which included gastrointestinal problems such as a peptic ulcer, cardiovascular problems such as high blood pressure and pregnancy problems such as premature birth and low birth weight. Most candidates achieved at least 1 mark, but some suggested things like 'lack of sleep' without clarifying what the effect of a lack of sleep on health would be.

(d) There were some very good answers in response to this question by candidates who related potential generalisations to the findings of the Gold et al. study. Only a few of these candidates provided an appropriate conclusion.

Section B

Question 5

(a) Many candidates chose to conduct an experiment rather than an observation and so were limited to L1 band marks. Candidates should always plan a study using the method stated in the question. Other candidates did an observation but needed to refer to the specific features of an observation, such as whether it was structured or unstructured, what response categories were or the number of observers, for example. Most answers also needed to address how data collected would be analysed to conclude whether a token economy effectively manages schizophrenia.

(b) The psychological evidence most apposite was the study by Paul and Lentz (1977) who studied schizophrenia using a token economy, but sometimes the token economy system used by Fox et al. in an open-cast mine was mentioned. Very few answers stated how this research had informed the planning of their study, as the question required. Methodologically most answers should have written more about the specific features of their observations. It is also worth a reminder that an evaluation of the design is inappropriate. Evaluation is not required in this question and there are no marks allocated to evaluation.

Question 6

(a) Very few answers designed appropriate studies in response to this question and most answers scored no more than L1 band marks. The question stated: '…experiment to investigate gender differences in pre-cognitive decisions' and was based on the syllabus section 'consumer decision-making', sub-section 'pre-cognitive decisions (Knutson et al., 2007)'. A pre-cognitive decision, as investigated by Knutson et al. is where the brain makes a decision before the person 'thinks it' and

before they speak it. This can only be achieved using an fMRI scanner. This is different from the studies designed by most candidates asking people what decision they make about a product.

(b) Very few answers mentioned the Knutson et al. (2007) study. This often restricted descriptions of psychological evidence to generalised comments and studies from other sections of the syllabus, such as 'thinking fast and thinking slow'. Methodological evidence was often inappropriate because it did not involve a fMRI scanner. The 'experiment' component was absent from many answers with questionnaires being used to ask people about products they had just purchased.

Question 7

(a) This question required candidates to plan a correlational study which meant that any data gathered would need to be plottable on a scatter graph. Whilst this was achieved in some plans, it was often absent in others. The simplest format would be to ask the parents to rate a child's pain using a visual analogue scale (for example) and then ask a medical practitioner to rate the same child's pain using the same scale. This could be repeated for perhaps twenty children and then the parent rating would be plotted on the x-axis and the practitioner rating on the y-axis.

(b) The visual analogue scale would be an appropriate measure to use for adults or perhaps the UAB pain rating scale (although that is usually for longer-term pain recording). The Wong-Baker scale faces was not really appropriate to use with adults (adult ratings were gathered, not the child's). Methodologically candidates often struggled to identify a specific method. Some wanted to make it an experiment, others a questionnaire or interview. The method used to gather data does not matter – the data just needs to be correlational.

Question 8

(a) There were two types of answer in response to this question, both equally acceptable. The first type involved an observation of errors with the machine in a real-life (field) setting with the person using the machine for their normal work. The second type involved bringing the machine into a laboratory and then manipulating one aspect to see what participants preferred (that would cause least errors/accidents). For example, one answer manipulated the size of a pressure dial, with the aim to see if the larger the dial the more it would be noticed and so acted upon more quickly and preventing an accident.

(b) Psychologically there were many features of a machine that could have been studied and these could include: controls (knobs, switches, levers, etc.) and displays (visual such as flashing lights and colours) and auditory (such as sirens and their volume). Importantly, human decision-making is also relevant such as the error of omission (failing to do something) because an alarm could be ignored if it is too quiet/cannot be heard, for example.

Section C

Question 9

Answers achieving the highest marks began with an outline of the genetic explanation of schizophrenia and then provided supporting research evidence which was usually the study by Gottesman and Shields (1972). The second half of the answer was then a discussion of how alternative explanations differed from the genetic explanation, in other words an evaluation of it. Such alternatives included the cognitive and biochemical explanations. Not all candidates took this approach and achieved lower marks, often describing a range of explanations without addressing the question.

Question 10

There were many answers which showed good understanding of the sub-topic and successfully evaluated the studies on brand recognition in children by Fischer et al. (1991) and product placement in films by Auty and Lewis (2004) which used children. There were many answers which stated 'children who are under 16 years cannot give consent' without addressing the point that consent can be given by their parent, teacher or appropriate adult. Many candidates appeared to consider that any study involving children is automatically unethical.

Question 11

The syllabus clearly distinguishes between three approaches to pain management: medical (biochemical), psychological and alternatives (acupuncture and stimulation therapy/TENS). This question provided the opportunity to present alternative techniques and then evaluate them using the other approaches. This was done very well by top-band answers, but at the bottom end of the mark range the approaches were often confused, particularly psychological and alternative. A few candidates considered medication, yoga and other strategies not on the syllabus, each of which received marks if they were appropriate to the question.

Question 12

The use of rating scales feature in every option of the syllabus because they are commonly used measures. Rating scales can be four-point, five-point or even seven-point or more. Candidates need to be able to consider the strengths and weaknesses of these types. Many answers provided a generalised answer of the strengths and weaknesses of questionnaires. Such strengths and weaknesses were not totally irrelevant and some marks were earned. The focus here should have been specifically on rating scales as applied to quality of working life, such as the Quality of Working Life (QWL) questionnaire by Walton (1974). This was rarely mentioned in answers.

PSYCHOLOGY

Paper 9990/42
Specialist Options: Application

Key messages

(a) What has been learned from the AS component of the syllabus should be transferred to the A2 component. For example, at AS candidates learn about methodology, such as experiments, which also apply to A2.

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

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

(d) In ***Section B***, **Questions 5**, **6**, **7** and **8**, methodological knowledge must be evident and detailed for top marks to be accessed. The procedure, however detailed is just one methodological aspect. For top marks answers must explain methodology rather than merely identify it.

(e) In ***Section C***, **Questions 9**, **10**, **11** and **12**, to access top marks, answers must include a debate which has two sides, such as strengths/advantages and weaknesses/disadvantages. Supporting evidence should also be provided. Description cannot be credited.

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

General comments

Some candidates appeared not to have studied two options at the same level of depth, and a few candidates answered questions from one option only. Some candidates answered questions from three and even four options. Whilst answers to one option were often very good, some answers to the second option were very poor, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

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

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

(iii) For **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so marks were limited.
Candidates should not use terms without explanation. Frequently answers stated 'it is reductionist' or 'it is useful in everyday life' without further explanation. To state 'it is reductionist' is merely to identify; it is not automatically a strength or weakness. Further many candidates assume that to be reductionist is always a weakness. It is not; any experiment is reductionist because variables are controlled and only the IV is manipulated. Reductionism is the basis of any experiment and as such it is a strength.

(iv) Many conclusions repeated what had already been written, and such summaries scored no marks. A conclusion is a 'decision reached by reasoning' and so as the reasoning has been done through the advantages and disadvantages, a final decision/conclusion needs to be drawn.

(v) Candidates should think about what the question requires rather than writing pre-prepared answers. Many questions will test the ability to apply knowledge from one thing to another, particularly methodological knowledge.

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

Section B

Many candidates conduct an experiment whatever the question. An interview, questionnaire or observation are methods independent of an experiment and candidates should not try to make other methods 'fit' into an experimental format. Answers to **part (a)** questions in this section should include an appropriate design, have applied a range (four or five) of relevant methodological design features, each of which should be explained fully, showing good understanding. Many answers listed features such as 'I would have a random sample' and 'It would be an independent measures design' without explanation of why it would be a random sample, or how this would be obtained.

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

Section C

It is essential that answers focus on the question that is set. Every question in this Section invites candidates to consider the extent to which they agree or disagree with the statement, rather than describe what they know about that topic area, and answers that do not address the question will only achieve minimal marks. To score marks at the top end of the mark range, answers must focus on arguments both for and against the statement, answers must the use appropriate evidence to support the argument, and, at the very top of the mark range, answers should show awareness of wider issues and evidence that is relevant.

Comments on specific questions

Section A

Question 1

(a) There were three types of answer in response to this question: (i) answers which suggested an appropriate physiological response and worded it in the form of a response statement such as 'my heartbeat speeds up', (ii) answers where there was a correct physiological response which needed to be worded as a response statement, for example 'heartbeat', and (iii) those who provided a cognitive response, for example 'I think I'm going to faint', rather than a physiological response, which could not be credited.

(b) This question was answered incorrectly by many candidates who wrote generally about anxiety scales rather than specifically 'using a four-point scale' as the question stated. Two strengths were required and each strength needed to be supported with an example related to people with blood injection phobia. The stem of the question, a four-point scale about blood phobia, was not addressed by most candidates.

(c) The question required one way to measure anxiety without using a rating scale, so any answer using a rating scale could not be credited. Many candidates suggested using a questionnaire that involved asking people open-ended questions and gathering qualitative data. It was unclear how anxiety could be measured in most of these responses. Many candidates suggested using the GAD 7, which uses a rating scale. Correct answers included suggestions such as observing the

behaviour of the person with blood anxiety and how they respond when they see a pool of blood on the floor.

(d) Many candidates wrote tautological responses, i.e. that one advantage of using quantitative data is that it is quantitative data, or stated 'it is quantitative' or that 'the data is numerical'. Candidates often wrote that an advantage is that the data is objective. However, the person providing the data may not be honest and so the data may be the subjective opinion of a person. This needed to be clarified for a mark to be awarded.

Question 2

(a) Most candidates were awarded 2 marks for stating that the two conditions of the independent variable were the quiet condition and the loud condition. Answers stating that the IV was the background noise were awarded 1 mark for a partially correct answer.

(b) Answers to this question could be any finding taken from the table of data appearing in the stem. Correct answers included: 'Sweetness was rated as being significantly lower in the loud condition than the quiet condition'. Alternatively, data could be used: 'Sweetness was rated as being significantly lower in the loud condition –3, than the quiet condition –0.2'.

(c) (i) This question required candidates to state the experimental design that was used by Woods et al. The experimental design used was repeated measures and identifying this design scored one mark. An additional mark was awarded for stating that in the Woods study participants completed both the quiet and loud conditions. Candidates stating that an independent groups design was used scored no marks as did candidates identifying the method (an experiment) which is very different from the design.

(ii) Correct answers suggested an independent design with a problem being that, for example, 'individual differences between participants might confound the result' and then relating this to the Woods et al. study 'e.g., one participant may prefer saltiness another not'.

(d) Most answers included two strengths and two weaknesses of using laboratory experiments, but these were often in general terms and needed to be focused on the question set. Answers needed to address the second part of the question, which was 'to investigate the effect of background noise on consumer behaviour'. Answers including reference to this often scored full marks. In most cases a summary was given rather than a conclusion.

Question 3

(a) Many candidates could not explain the term 'pain'. Answers like 'it is a sensory experience' were too vague. Answers also referred to pain being a physiological and psychological event. To be awarded full marks, answers needed to provide some detail or give an example. Referring to acute or chronic pain, or phantom limb pain would have been sufficient. One good explanation of pain is that it is 'an unpleasant sensory and emotional experience (sensory and/or emotional discomfort) associated with actual or potential tissue damage, or described in terms of tissue damage, or both'.

(b) Most candidates were awarded full marks for providing two alternative techniques, both identified and with some detail/elaboration. Acupuncture and transcutaneous electrical nerve stimulation (TENS) were often very clearly described. A few candidates suggested meditation or yoga. Some suggested psychological techniques which did not address the question set.

(c) There are three psychological techniques: attention diversion, non-pain imagery and cognitive redefinition and nearly all candidates opted to outline one of these three. The technique was often identified but needed some detail/elaboration.

(d) Although there was some confusion over the term 'medical techniques', most candidates provided two strengths and two weaknesses. Most candidates did not relate their answers to the question – in this instance pain, for example 'it is easy to take a pill' or 'medicines may have side effects', which needed to be linked to pain to be creditable, for example 'it is easy to take a pill, such as taking paracetamol (or equivalent) for a painful headache'.

Question 4

(a) Nearly all candidates scored limited credit for 'followers are people who follow the leader' but more was required for full marks. For example, using the information in the stem, answers could have referred to the different types of follower, or provided an example of one type such as 'effective followers'.

(b) There were some strong answers to this question, but some were not able to show understanding of Kelly's (1988) work. Some focussed on different types of followers and needed to focus on qualities.

(c) There were some strong answers, often providing clear and simple but correct differences. For example, 'conformist followers do not question the leader whereas alienated followers always question the leader.' Some answers confused conformist followers and sometimes alienated followers with other types. Responses needed to focus on differences, rather than points of description in isolation.

(d) Some excellent responses were written to this question, using Kelly's types to enhance their points.

Section B

Question 5

(a) Candidates were invited to design a study with no stated method meaning that any method could be chosen. Most opted for an experiment comparing SSRIs with a control group. Most candidates included the specific features of this method, but often these features were identified rather than explained. For example, 'an independent design was used' and 'randomised control trials were used' without further elaboration. The same was often true for general design features, with a sampling technique being identified with a statement 'I would use opportunity sampling', which needed a little more explanation to enhance the quality of the answer.

(b) The psychological evidence often included descriptions of how SSRIs work but would have benefitted from addressing how long SSRIs need to be taken for before they become effective, or what a typical dosage would be. Some responses demonstrated limited knowledge of OCD, for instance stating that people with OCD live in a hospital or mental institution. In most cases people live at home and follow a relatively normal life. Methodologically, there was limited description on how the effectiveness of the SSRIs would be measured, with most candidates stating 'the effectiveness would be measured' but needed to explain how this would be done. A number of candidates evaluated their design, which is not required by the question. An explanation was needed of why certain design features were used in preference to others.

Question 6

(a) This question required the use of a questionnaire. Some responses conducted an experiment where people ordered drinks and then gave a questionnaire asking about choice of drinks, when the choice was already known. Questionnaires can be conducted in a restaurant, but they can be conducted anywhere. There was also confusion about the term primacy and recency, sometimes leading to confused designs.

(b) Many candidates wrote about the work of Dayan and Bar-Hillel (2011) and gave a strong description of what they did there but needed to explain how their work informed this study. There was also confusion over primacy and recency. Methodologically, candidates should choose at least two design features (specific or general) and explain why they chose those particular features rather than any other, such as why closed questions were used rather than open-ended.

Question 7

(a) There were some excellent answers who answered the question set clearly and unambiguously. The question stated 'Design a case study…' but most responses did not use the case study method, resulting in L1 band marks. A case study is usually conducted on one person, or one 'unit' such as a school. Opting to do a case study on a school would be acceptable, provided this was stated. A second common problem was that the providing information strategy should have been used (as stated in the question) but many answers focused instead on fear arousal, and credit was limited.

(b) In relation to psychological evidence, as many responses focused on fear arousal they described the work of Janis and Feshbach which was not irrelevant to this question. The work of Lewin (1992) on providing information (he designed the Heart Health Manual) was apposite, and some candidates wrote about Tapper et al.'s food dudes and this could be made relevant. Methodologically answers were often irrelevant because of an incorrect method, such as an experiment, being conducted in response to **part (a)**. Candidates must always focus on the question set.

Question 8

(a) Candidates were required to conduct an observation, and although there were a few excellent answers, most candidates needed to focus more on the specific features of an observation. For example, was their design covert or overt; participant or non-participant; controlled or naturalistic; structured or unstructured.

(b) Psychologically, most candidates scored some marks for their knowledge of forming and norming, although a few answers tended to be essay-like and included full details of all Tuckman's stages, which was not required. Methodologically, answers were often sparse due to the lack of specific features about observations in **part (a)**. To improve, for example, responses could explain why the observation was structured rather than unstructured, and allow the different features of forming and norming to be identified and observed.

Section C

Question 9

Answers achieving the highest marks began with an outline of the cognitive approach in relation to phobias then provided supporting research evidence which was usually the study by DiNardo on dogs. The second half of the answer was then a discussion of how alternative explanations differed from the cognitive explanation. Such alternatives included the behavioural, psychoanalytic and biochemical explanations. Not all candidates took this approach and achieved lower marks. Some answers focused on the cognitive approach in general without relating to phobias. Others described the cognitive and other explanations and needed to address the question set. A small number of candidates focused on cognitive treatments for depression which could not be credited.

Question 10

At the top end of the mark range the study by Gil et al. was thoroughly considered with relevant arguments both for and against movement patterns presented. A few answers focused exclusively on store interior layout, whereas top band answers used this knowledge to extend their discussion of movement patterns. There were some anecdotal responses about candidates' own shopping experiences which needed to be based on psychological theory and research in order to be creditable.

Question 11

The Yale model of communication did not appear to be understood by many candidates. Many answers were a holism versus reductionism debate, with brief reference to the 'Yale model' sometimes included. There appeared to be a common assumption that reductionism is always negative, it wasn't always understood that reductionism is the basis of any experiment where the IV is isolated and all other variables are controlled – what is studied is one component taken from the whole behaviour (or whatever is being studied).

Question 12

The focus of this question was on the Leadership Practices Inventory (LPI) and the strengths and weaknesses of the LPI-self and LPI-observer questionnaires. There were some excellent answers, but others provided a generalised answer of the strengths and weaknesses of questionnaires that not focussed on the context of the question. Such strengths and weaknesses were not irrelevant, and some marks were earned. Responses need be able to apply such strengths and weaknesses to gain more credit.

PSYCHOLOGY

Paper 9990/43
Specialist Options: Application

Key messages

(a) What has been learned from the AS component of the syllabus should be transferred to the A2 component. For example, at AS candidates learn about methodology, such as experiments, which also apply to A2.

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

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

(d) In ***Section B***, **Questions 5**, **6**, **7** and **8**, methodological knowledge must be evident and detailed for top marks to be accessed. The procedure, however detailed is just one methodological aspect. For top marks answers must explain methodology rather than merely identify it.

(e) In ***Section C***, **Questions 9**, **10**, **11** and **12**, to access top marks, answers must include a debate which has two sides, such as strengths/advantages and weaknesses/disadvantages. Supporting evidence should also be provided. Description cannot be credited.

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

General comments

Some candidates appeared not to have studied two options at the same level of depth, and a few candidates answered questions from one option only. Some candidates answered questions from three and even four options. Whilst answers to one option were often very good, some answers to the second option were very poor, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

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

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

(iii) For **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so marks were limited.
Candidates should not use terms without explanation. Frequently answers stated 'it is reductionist' or 'it is useful in everyday life' without further explanation. To state 'it is reductionist' is merely to identify; it is not automatically a strength or weakness. Further many candidates assume that to be reductionist is always a weakness. It is not; any experiment is reductionist because variables are controlled and only the IV is manipulated. Reductionism is the basis of any experiment and as such it is a strength.

(iv) Many conclusions repeated what had already been written, and such summaries scored no marks. A conclusion is a 'decision reached by reasoning' and so as the reasoning has been done through the advantages and disadvantages, a final decision/conclusion needs to be drawn.

(v) Candidates should think about what the question requires rather than writing pre-prepared answers. Many questions will test the ability to apply knowledge from one thing to another, particularly methodological knowledge.

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

Section B

Many candidates conduct an experiment whatever the question. An interview, questionnaire or observation are methods independent of an experiment and candidates should not try to make other methods 'fit' into an experimental format. Answers to **part (a)** questions in this section should include an appropriate design, have applied a range (four or five) of relevant methodological design features, each of which should be explained fully, showing good understanding. Many answers listed features such as 'I would have a random sample' and 'It would be an independent measures design' without explanation of why it would be a random sample, or how this would be obtained.

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

Section C

It is essential that answers focus on the question that is set. Every question in this Section invites candidates to consider the extent to which they agree or disagree with the statement, rather than describe what they know about that topic area, and answers that do not address the question will only achieve minimal marks. To score marks at the top end of the mark range, answers must focus on arguments both for and against the statement, answers must the use appropriate evidence to support the argument, and, at the very top of the mark range, answers should show awareness of wider issues and evidence that is relevant.

Comments on specific questions

Section A

Question 1

(a) This question asked for a behavioural explanation of phobias, and nearly all answers were awarded 1 mark for stating that 'all behaviours are learned'. Some answers went further and explained how this was done and some answers, also awarded 2 marks, gave the example of little Albert who was conditioned to be afraid of a white rat.

(b) There were some excellent answers in response to this question, but there were also some zero mark answers. The strongest answers included the two essential elements: the creation of an anxiety hierarchy and the learning of progressive muscle relaxation. Inclusion of these two elements scored 2 marks and for the 'in this study' component, relating Richard's fear of oranges to systematic desensitisation, scored the additional 2 marks.

(c) Two other ways to treat a phobia that are included on the syllabus are Ost's applied tension (to treat blood/injection phobia) and cognitive behaviour therapy (Ost and Westling) which can be used to treat any phobia. Many candidates opted for these two, but other possibilities were also acceptable in this 'suggest' question. For example, psychoanalytic psychotherapy was sometimes mentioned.

(d) When considering the strengths and weaknesses of using systematic desensitisation most candidates provided logical and appropriate responses. However, a few answers needed to acknowledge that people undergoing systematic desensitisation are not participants in a study but real people with real phobias. Comments such as 'participants might try to move through the hierarchy quicker' are incorrect; the opposite is often the case. Similarly, 'the participant might be harmed by the treatment, so it is unethical' is also incorrect because the person has chosen the treatment to reduce the effects of the phobia and also because the hierarchy is designed so it is not distressing or harmful.

Question 2

(a) There were three types of answer in response to this question: those who had no understanding of the retroactive effect (or interference); those who confused it with proactive interference; and those who knew and understood the term. Those who related retroactive interference to the Braun-LaTour study were awarded 2 marks.

(b) (i) Many candidates were not able to give an example of the data collected, and are advised to carefully read the stem of the question which prompts the response.

(ii) Most candidates were awarded full marks. In this instance there was no requirement to relate it to the study, although this was creditable as explanation.

(c) The situational component would involve the past experience of visiting Disneyland and whether they had seen certain characters there, such as Bugs Bunny. The individual component involved the false memory (a cognitive process) that was created in the mind of the participant. Some excellent answers focused clearly on these two and showed good understanding.

(d) Many answers included two strengths and two weaknesses of giving course credits to students, but often answers only scored partial marks because answers were not related to the study of false advertising as the question required. A simple formula applies to all **part (d)** questions: strength + example, strength + example, weakness + example, weakness + example. Note that **(i)** 1 mark required strength or weakness *plus* an example; **(ii)** the question required strengths and weaknesses (plural) not just one of each.

Question 3

(a) Most candidates could not outline an appropriate conclusion from the Riekert and Drotar study. Two conclusions appear in the abstract of the Riekert and Drotar study and either could be credited.

(b) (i) Many suggested that the children were under 16 years and could not give consent. Notably in any study an adult may give consent for a child when the child does not wish it. In this study the researchers required consent from *both* parent and child. Correct answers were therefore that 14 of the families (both adolescent and parent) did not give consent to participate. Another correct answer reducing the number of participants was that 28 participants (labelled 'non-returners') did not return the postal questionnaire as they applied their right to withdraw.

(ii) There was a wide range of incorrect answers to this question. Excluding participants by researchers from any study is important in psychological research because it might be that participants might be harmed by the research, for example. Given the nature of the study, Riekert and Drotar excluded participants because: they had been diagnosed with insulin-dependent diabetes mellitus or because they had an additional chronic illness or because they showed a developmental delay such as Down's Syndrome.

(c) Most answers were awarded maximum marks for explaining two reasons why a patient might not adhere to medical advice. Rational non-adherence was given by most candidates and factors from the health belief model were also appropriate. A few candidates provided 'the cost of the medicine'

which was awarded 1 mark because it is not incorrect but needed more detail for it to be awarded another mark.

(d) There were many answers related to the strengths and weaknesses of questionnaires in general, such as 'people might not give honest answers' but this question was about 'using questionnaires to be returned by post' which is rather more specific. Candidates should always read the question carefully and address the focus in their answer. Some answers did address the specific question but needed to relate their answers to gathering data about health as the question required.

Question 4

(a) Most candidates correctly outlined a 'rotator', writing 'a worker who works a pattern of rotating (alternating) shifts of days/evenings/nights' and adding, for example 'compared to day workers, evening workers, night workers and day workers with occasional nights' for the extra detail.

(b) Nearly all candidates scored 1 mark for one finding, but the question required two findings. Some answers were correct but needed to have more detail. Writing 'night workers had less than 5 hours sleep' is correct and scored 1 mark. Writing '20% of night workers had less than 5 hours sleep compared to other groups such as rotators' scored 2 marks.

(c) (i) This question required candidates to identify two types of error or accident more likely to happen to nurses on a rotator shift. Some candidates did not address the latter part of the question and gave two generalised responses. Others successfully identified two of the four possibilities.

(ii) This question required candidates to suggest two effects shiftwork can have on health, and they could have used the findings of the Knuttson (2003) study, which included gastrointestinal problems such as a peptic ulcer, cardiovascular problems such as high blood pressure and pregnancy problems such as premature birth and low birth weight. Most candidates achieved at least 1 mark, but some suggested things like 'lack of sleep' without clarifying what the effect of a lack of sleep on health would be.

(d) There were some very good answers in response to this question by candidates who related potential generalisations to the findings of the Gold et al. study. Only a few of these candidates provided an appropriate conclusion.

Section B

Question 5

(a) Many candidates chose to conduct an experiment rather than an observation and so were limited to L1 band marks. Candidates should always plan a study using the method stated in the question. Other candidates did an observation but needed to refer to the specific features of an observation, such as whether it was structured or unstructured, what response categories were or the number of observers, for example. Most answers also needed to address how data collected would be analysed to conclude whether a token economy effectively manages schizophrenia.

(b) The psychological evidence most apposite was the study by Paul and Lentz (1977) who studied schizophrenia using a token economy, but sometimes the token economy system used by Fox et al. in an open-cast mine was mentioned. Very few answers stated how this research had informed the planning of their study, as the question required. Methodologically most answers should have written more about the specific features of their observations. It is also worth a reminder that an evaluation of the design is inappropriate. Evaluation is not required in this question and there are no marks allocated to evaluation.

Question 6

(a) Very few answers designed appropriate studies in response to this question and most answers scored no more than L1 band marks. The question stated: '…experiment to investigate gender differences in pre-cognitive decisions' and was based on the syllabus section 'consumer decision-making', sub-section 'pre-cognitive decisions (Knutson et al., 2007)'. A pre-cognitive decision, as investigated by Knutson et al. is where the brain makes a decision before the person 'thinks it' and

before they speak it. This can only be achieved using an fMRI scanner. This is different from the studies designed by most candidates asking people what decision they make about a product.

(b) Very few answers mentioned the Knutson et al. (2007) study. This often restricted descriptions of psychological evidence to generalised comments and studies from other sections of the syllabus, such as 'thinking fast and thinking slow'. Methodological evidence was often inappropriate because it did not involve a fMRI scanner. The 'experiment' component was absent from many answers with questionnaires being used to ask people about products they had just purchased.

Question 7

(a) This question required candidates to plan a correlational study which meant that any data gathered would need to be plottable on a scatter graph. Whilst this was achieved in some plans, it was often absent in others. The simplest format would be to ask the parents to rate a child's pain using a visual analogue scale (for example) and then ask a medical practitioner to rate the same child's pain using the same scale. This could be repeated for perhaps twenty children and then the parent rating would be plotted on the x-axis and the practitioner rating on the y-axis.

(b) The visual analogue scale would be an appropriate measure to use for adults or perhaps the UAB pain rating scale (although that is usually for longer-term pain recording). The Wong-Baker scale faces was not really appropriate to use with adults (adult ratings were gathered, not the child's). Methodologically candidates often struggled to identify a specific method. Some wanted to make it an experiment, others a questionnaire or interview. The method used to gather data does not matter – the data just needs to be correlational.

Question 8

(a) There were two types of answer in response to this question, both equally acceptable. The first type involved an observation of errors with the machine in a real-life (field) setting with the person using the machine for their normal work. The second type involved bringing the machine into a laboratory and then manipulating one aspect to see what participants preferred (that would cause least errors/accidents). For example, one answer manipulated the size of a pressure dial, with the aim to see if the larger the dial the more it would be noticed and so acted upon more quickly and preventing an accident.

(b) Psychologically there were many features of a machine that could have been studied and these could include: controls (knobs, switches, levers, etc.) and displays (visual such as flashing lights and colours) and auditory (such as sirens and their volume). Importantly, human decision-making is also relevant such as the error of omission (failing to do something) because an alarm could be ignored if it is too quiet/cannot be heard, for example.

Section C

Question 9

Answers achieving the highest marks began with an outline of the genetic explanation of schizophrenia and then provided supporting research evidence which was usually the study by Gottesman and Shields (1972). The second half of the answer was then a discussion of how alternative explanations differed from the genetic explanation, in other words an evaluation of it. Such alternatives included the cognitive and biochemical explanations. Not all candidates took this approach and achieved lower marks, often describing a range of explanations without addressing the question.

Question 10

There were many answers which showed good understanding of the sub-topic and successfully evaluated the studies on brand recognition in children by Fischer et al. (1991) and product placement in films by Auty and Lewis (2004) which used children. There were many answers which stated 'children who are under 16 years cannot give consent' without addressing the point that consent can be given by their parent, teacher or appropriate adult. Many candidates appeared to consider that any study involving children is automatically unethical.

Question 11

The syllabus clearly distinguishes between three approaches to pain management: medical (biochemical), psychological and alternatives (acupuncture and stimulation therapy/TENS). This question provided the opportunity to present alternative techniques and then evaluate them using the other approaches. This was done very well by top-band answers, but at the bottom end of the mark range the approaches were often confused, particularly psychological and alternative. A few candidates considered medication, yoga and other strategies not on the syllabus, each of which received marks if they were appropriate to the question.

Question 12

The use of rating scales feature in every option of the syllabus because they are commonly used measures. Rating scales can be four-point, five-point or even seven-point or more. Candidates need to be able to consider the strengths and weaknesses of these types. Many answers provided a generalised answer of the strengths and weaknesses of questionnaires. Such strengths and weaknesses were not totally irrelevant and some marks were earned. The focus here should have been specifically on rating scales as applied to quality of working life, such as the Quality of Working Life (QWL) questionnaire by Walton (1974). This was rarely mentioned in answers.