PSYCHOLOGY

Paper 9990/11
Approaches, Issues and debates

Key messages

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

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

Candidates need to be careful about how they present 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 need to appreciate the difference between features of a sample and characteristics of a sample. Characteristics are the participant variables that are presented by the sample, whereas features are anything related to the sample, for example, characteristics, sample size, sampling technique etc.

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

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

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

Candidates 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, which 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.

General comments

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

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

Comments on specific questions

Question 1

(a) Some responses could identify the correct sample size used in the study by Bandura et al. Other responses provided sample sizes for other core studies.

(b) Stronger responses could clearly describe the pre-study assessment of aggression in the children. Many responses focused, incorrectly, on the measurement of aggression in the final part of the study. It is important for candidates to read the question carefully to ensure that they are responding with the correct part of the study.

(c) Candidates need to appreciate the difference between a conclusion and a result. The former is a generic comment based on the results reported in any core study. The latter is factual and based on collected data. Therefore, to gain credit on this question candidates needed to give a generic comment based around aspects like social learning or modelling. A significant minority of responses gave a factual result based on imitative results from the study but could not gain credit here.

Question 2

(a) For these types of questions, responses should focus on the general psychology that is being investigated in the study rather than a specific aim of the study. Therefore, aspects of the Andrade study that could gain credit here included dual processing, boredom, and attention. Credit could be given to generic descriptions of the principles of either of these. However, many responses focused too narrowly on the aims of Andrade and what she found in the study. These could only gain one available example mark.

(b) Stronger responses covered both the identification of a strength linked to the experimental design and then context from the study by Andrade. Popular choices included reduction in demand characteristics and the potential lack of order effects. Some responses focused on the strengths of using an experiment or incorrectly presented the experimental design as repeated measures.

Question 3

(a) Stronger responses could clearly outline the sampling technique used by Canli et al. Some responses could identify volunteer sampling as the technique but then did not gain the second available mark as they then used the word to define the word, for example, by stating ‘this meant participants volunteered.’ This does not explicitly demonstrate understanding so cannot gain credit. To improve, candidates need to be able to define all key terms without using that term in the definition.

(b) Some responses could clearly outline the two questions asked from participants during the recognition test. However, many responses tended to focus on the ratings that participants had to give to the images, but this was from a different part of the study so could not gain credit. It is important for candidates to know the full procedure of every Core Study and to separate out each part of those procedures.

(c) Some responses could describe the positive correlation reported by Canli et al. However, some responses gave a different second measure in their explanation (e.g. amygdala arousal) which was not part of the question. These responses could only gain the mark available for identifying the type of correlation. Other responses did not demonstrate knowledge of a what a correlation is and described other results or separate results for the two measures in the question.

Question 4

(a) Stronger responses could clearly describe what the victim was expected to do during any trial. Common descriptions were them falling to the ground in the critical area and remaining there until help was given. However, many responses focused on the difference between the two types of victim which could not gain credit as it was not answering the question set. It is important for candidates to know all aspects of every Core Study on the syllabus to get a comprehensive

overview of each one. It is also important to read the question carefully to ensure that a response is based on that part of the procedure.

(b) Many responses could identify an appropriate weakness. These could be awarded the available mark by explicitly stating what the weakness is. Therefore, stating 'it lacked generalisability' could gain the mark but simply stating 'generalisability' by itself could not, as the candidate had not identified what the weakness is. Other creditworthy responses focused on lack of controls or breaking ethical guidelines. It is important for candidates to be explicit in their responses to ensure that they are answering the question set and so the Examiner does none of the work. To improve on these types of questions, candidates need to be explicit in their responses as to why something is a strength/weakness.

Question 5

(a) A minority of responses could identify two characteristics of the sample. Common choices were the majority being female, and students. A minority of responses gave features from samples used in other Core Studies rather than that of Laney et al. or features of Laney et al. like the sampling technique that could not gain credit. To improve, candidates need to appreciate the difference between features of a sample and characteristics of a sample. Characteristics are the participant variables that are presented by the sample, whereas features are anything related to the sample, for example, characteristics, sample size, sampling technique etc.

(b) Stronger responses could clearly outline a potential problem and then put it into the context of the study by Laney et al. Common responses included it being fixed choice, and not predicting what a person may actually pay in reality. Many responses could outline a generic research methods based problem with the question about food costs but then did not give an example based on the study itself, only gaining partial credit. To improve, candidates need to remember to explicitly link their responses to the study named in the question to be able to access all available marks.

Question 6

(a) Weaker responses tended to describe what Alex the parrot achieved in the study which could only gain minimal credit. Stronger responses could clearly describe many aspects of Social Learning Theory including observation, imitation, replication, and motivation. To improve, candidates need to be able to describe key ideas and theories linked to each Core Study in isolation from the actual Core Study itself.

(b) For this type of question, responses must contain two parts. The first is an appropriate result. The second is an explicit link to the concept named in the question, in this case, social learning theory. Some responses did present an appropriate result but did not then explicitly explain why the result could support social learning theory. Stronger responses demonstrated both parts of the question and could clearly link back to social learning theory via the model/rival technique.

Question 7

(a) Many responses could clearly outline what was meant by informed consent. However, there was a significant minority of responses that gave tautological answers, and these could not be credited. For example, a response may correctly give the 'when the participant is given enough information' part of the ethical guideline but finish with 'and therefore the participant can give their consent'. Using the word to define the word cannot gain credit in an examination as it does not demonstrate that the candidate understands the full term. To improve, candidates need to be able to define key terminology without using the term itself.

(b) The majority of responses could identify one ethical guideline that was not broken in the study by Dement and Kleitman. Stronger responses then used evidence directly from the study to outline why the guideline(s) had been broken. Popular choices were confidentiality, privacy and informed consent. Weaker responses tended to simply describe ethical guidelines without using evidence to show how it was broken. To improve, candidates need to read the question carefully as in this instance, the focus was on using evidence from the study.

Question 8

(a) Some responses could clearly outline the individual and situational explanations used in psychology. However, a significant number of these, like with **Question 7(a)**, used the term to define the term. A common example was stating that the situational explanation is based around the situation a person finds themselves in. This cannot gain credit as it does not demonstrate understanding. Stronger responses could give clear examples from studies other than Milgram including differences in participants' doodling (individual: Andrade) and differences in aggression after watching the model (situational: Bandura). Some responses used Milgram as the example and could not gain credit. To improve, like with **Question 7(a)**, candidates need to be able to define key terminology without using the term itself.

(b) Candidates need to appreciate the difference between a result and a conclusion. The former is factual and based on collected data. The latter is a generic comment based on the results reported in any core study. This question was asking about both of these. Weaker responses could give either a brief result or a brief conclusion from the study by Milgram and nothing else. However, there were some stronger responses that could demonstrate knowledge of a conclusion from the study and then justify this with the data collected in the study. Popular choices included that people will follow destructive orders, and that people show overt signs of distress when being obedient.

Question 9

The strongest responses evaluated the study by Baron-Cohen et al. in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of generalisations. Common choices included types of data collected, reliability, validity, generalisability, and ethics. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Baron-Cohen et al. to explicitly support their point and 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 Baron-Cohen et al. as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. In addition, some candidates followed a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). Therefore, some responses appeared to be prepared essays for Baron-Cohen et al. A response that does not have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context, which are Level 2 responses. Any response that has no context cannot get above a Level 1 mark. In addition, some responses did use generalisability in an evaluative sense but did not fully explain <u>why</u> it could be a strength and/or a weakness. Stronger responses could identify the potential weakness of the demographics of any of the four groups and then give more than one example from the study making it 'in detail'. Some responses did not cover the named issue, only describing what generalisations mean or describing the different groups which could not gain credit. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue. Each strength and weakness should be of equal length with an explanation as to <u>why</u> it is a strength or weakness with examples from the study to demonstrate clear understanding. These are the requirements for a Level 4 response.

PSYCHOLOGY

Paper 9990/12
Approaches, issues and debates

Key messages

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

Candidates need to read the whole question carefully to ensure that their responses are fulfilling the demands of each one. For example, the question may require data, a named issue to be included or relate back to a previous answer. To achieve full marks, these need to be correctly present in their responses. The essay (final question) requires four evaluation points to be 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 present 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 need to appreciate the difference between features of a sample and characteristics of a sample. Characteristics are the participant variables that are presented by the sample, whereas features are anything related to the sample, for example, characteristics, sample size, sampling technique etc.

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

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

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

Candidates 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, which 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.

General comments

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

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

Comments on specific questions

Question 1

(a) The majority of responses could identify his mother as the cause of positive reinforcement. Incorrect responses included a therapist and Saavedra.

(b) Stronger responses could clearly name two of the stimuli fully. Many responses either simply stated 'button' or only part of the stimuli used in the fear hierarchy. Credit could only be given to full naming of the five types of buttons.

(c) Candidates need to appreciate the difference between a conclusion and a result. The former is a generic comment based on the results reported in any core study. The latter is factual and based on collected data. Therefore, to gain credit on this question, candidates needed to give a generic comment based around helping the boy with his phobia. A significant minority of responses gave a factual result based on distress scores but could not gain credit here.

Question 2

(a) For these types of questions, responses should focus on the general psychology that is being investigated in the study rather than a specific aim of the study. Therefore, aspects of the Pepperberg study that could gain credit here included social learning theory and operant conditioning. Credit could be given to generic descriptions of the principles of either of these. However, many responses focused too narrowly on the aims of Pepperberg and what she achieved with Alex the parrot. These could only gain one available example mark.

(b) Stronger responses covered both the identification of a strength linked to validity and then context from the study by Pepperberg. Popular choices included the reduction of researcher bias and the use of controls for internal validity. Some responses mixed controls with standardisation (reliability) so could not gain credit here.

Question 3

(a) There were very few correct responses to this question. Stronger responses could clearly outline two reasons for a relationship *before* the study by Canli et al. The majority of responses gave explanations about the findings *from* the actual study by Canli et al., but the question was not asking for this. It is very important that candidates know the background of each core study found in the introduction of each journal paper to ensure they understand the process involved in each one from beginning to end. This includes knowing some of the reasons why a study was conducted.

(b) Some responses could describe the negative correlation reported by Canli et al. due to the way the scale was measure for valence. It is important that candidates understand how a measure is scored in any core study as it may go against a logical format. This is the case here. The majority of responses reported it as a positive correlation which, because of how valence was scored, is incorrect.

Question 4

(a) Stronger responses could clearly describe the familiarisation phase of the study. Common descriptions included being able to freely manipulate tools and how long each phase lasted for. However, the majority of responses described either the Can See or the Cannot See condition. It is important for candidates to know all aspects of every Core Study on the syllabus to get a comprehensive overview of each one.

(b) Stronger responses could fully outline both possible outcomes that ended a trial. Many responses could only partially outline an outcome so could not gain credit. For example, one possible outcome was when the recipient chimpanzee received the correct tool and obtained the juice reward. Outlining the first part only could not gain credit as this did not mean a trial had ended. Again, candidates need to know how all aspects of every Core Study were measured and operationalised.

(c) Many responses could identify an appropriate methodological weakness. These could be awarded the available mark by explicitly stating what the weakness is. Therefore, stating 'it lacked ecological validity' could gain the mark but simply stating 'ecological validity' by itself could not as the candidate had not identified what the weakness is. Other creditworthy responses focused on generalisability and repeated measures. It is important for candidates to be explicit in their responses to ensure that they are answering the question set and so the Examiner does none of the work. To improve on these types of questions, candidates need to be explicit in their responses as to why something is a strength/weakness.

Question 5

(a) The majority of responses could identify two features of the sample. Common choices were being volunteers, the sample size and that they were all male. A minority of responses gave features from samples used in other Core Studies rather than that of Schachter and Singer.

(b) Stronger responses could clearly outline a potential problem and then put it into the context of the study by Schachter and Singer. Common responses included it being subjective and not predicting what a person is actually feeling. Many responses could outline a generic research methods based problem with the question about mood but then did not give an example based on the study itself, only gaining partial credit. To improve, candidates need to explicitly link their responses to the study named in the question to be able to access all available marks.

Question 6

(a) Weaker responses tended to describe the eyes test and how it works which could only gain minimal credit. Stronger responses could clearly describe many aspects of Theory of Mind including attribution of mental states and aspects of empathy. To improve, candidates need to be able to describe key ideas and theories linked to each Core Study in isolation from the actual Core Study itself.

(b) For this type of question, responses must contain two parts. The first is the 'what' – what real-world application could be. The second is the 'how' – how will the real-world application be achieved in an ethical way. Common responses included using the eyes test as a diagnostic tool and using the eyes test to help people understand emotions better in the classroom.

Question 7

(a) Many responses could clearly outline what was meant by informed consent. However, there was a significant minority of responses that gave tautological answers, and these could not be credited. For example, a response may correctly give the 'when the participant is given enough information' part of the ethical guideline but finish with 'and therefore the participant can give their consent'. Using the word to define the word cannot gain credit in an examination as it does not demonstrate that the candidate understands the full term. To improve, candidates need to be able to define key terminology without using the term itself.

(b) The majority of responses could identify one ethical guideline that was broken in the study by Milgram. Stronger responses then used evidence directly from the study to outline why the guideline(s) had been broken. Popular choices were deception, psychological stress, and the right to withdraw. Weaker responses tended to simply describe ethical guidelines without using evidence to show how it was broken. To improve, candidates need to read the question carefully as in this instance, the focus was on using evidence from the study.

Question 8

(a) Some responses could clearly outline the individual and situational explanations used in psychology. However, a significant number of these, like with **Question 7(a)**, used the term to define the term. A common example was stating that the situational explanation is based around the situation a person finds themselves in. This cannot gain credit as it does not demonstrate understanding. Stronger responses could give clear examples from the study by Dement and Kleitman with different dream content (individual) and being in a laboratory (situational) being popular choices. To improve, like with **Question 7(a)**, candidates need to be able to define key terminology without using the term itself.

(b) Candidates need to appreciate the difference between a result and a conclusion. The former is factual and based on collected data. The latter is a generic comment based on the results reported in any core study. This question was asking about both of these. Weaker responses could give either a brief result or a brief conclusion from the study by Dement and Kleitman and nothing else. However, there were some stronger responses that could demonstrate knowledge of a conclusion from the study and then justify this with the data collected in the study. Popular choices included people tending to dream more in REM and that eye movements in dreams could be linked to content of the dream.

Question 9

The strongest responses evaluated the study by Bandura et al. in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of quantitative data. Common choices included types of data collected, reliability, validity, generalisability, and ethics. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Bandura et al. to explicitly support their point and tended to score Level 4 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses, all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by Bandura et al. as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. In addition, some candidates followed a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). Therefore, some responses appeared to be prepared essays for Bandura et al. without one of their points being about quantitative data. A response that does not have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did use quantitative in an evaluative sense but did not fully explain why it could be a strength and/or a weakness. Stronger responses could identify the potential strength of ease of comparison or objectivity and then give more than one example from the study making it 'in detail'. Several responses did not cover the named issue, only describing what quantitative data is or incorrectly evaluated the named issue by giving strengths/weaknesses of qualitative data. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples from the study to show clear understanding. These are the requirements for a Level 4 response.

PSYCHOLOGY

Paper 9990/13
Approaches, Issues and Debates

Key messages

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

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

Candidates need to be careful about how they present 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 need to appreciate the difference between features of a sample and characteristics of a sample. Characteristics are the participant variables that are presented by the sample, whereas features are anything related to the sample, for example, characteristics, sample size, sampling technique etc.

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

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

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

Candidates 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, which 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.

General comments

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

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

Comments on specific questions

Question 1

(a) Some responses could identify the correct sample size used in the study by Andrade. Other responses provided sample sizes for other core studies.

(b) Stronger responses could clearly identify two features of the mock telephone message. Some responses focused, incorrectly, on the doodling aspect of the procedure. It is important for candidates to read the question carefully to ensure that they are responding with the correct part of the study.

(c) Candidates need to appreciate the difference between a conclusion and a result. The former is a generic comment based on the results reported in any core study. The latter is factual and based on collected data. Therefore, to gain credit on this question, candidates needed to give a generic comment based around aspects like dual-processing and how doodling can focus attention. A significant minority of responses gave a factual result based on recall of results from the study but could not gain credit here.

Question 2

(a) For these types of questions, responses should focus on the general psychology that is being investigated in the study rather than a specific aim of the study. Therefore, aspects of the Yamamoto et al. study that could gain credit here included empathy, altruism and targeting helping. Credit could be given to generic descriptions of the principles of either of these. However, some responses focused too narrowly on the aims of Yamamoto et al. and what they found in the study. These could only gain one available example mark.

(b) Stronger responses covered both the identification of a strength linked to validity and then context from the study by Yamamoto et al. Popular choices included cause and effect, and the reduction of extraneous variables. Some responses focused on the strengths of using an experiment or incorrectly debated standardisation and reliability.

Question 3

(a) Stronger responses could clearly outline the sampling technique used by Canli et al. Some responses could identify volunteer sampling as the technique but did not gain the second available mark as they then used the word to define the word, for example, by stating ‘this meant participants volunteered.’ This does not explicitly demonstrate understanding so cannot gain credit. To improve, candidates need to be able to define all key terms without using that term in the definition.

(b) The majority of responses scored at least one mark for identifying one of the characteristics of the victim (learner). Popular choices included being male and that he was mild-mannered. There were a significant minority of responses that reported characteristics of the sample used in the study by Milgram or of the experimenter who gave the prods. It is important that candidates clearly know the role of all people involved in the set-up of this particular study.

(c) Some responses could describe two changes to the shock generator. Popular choices were an electrical buzzing noise and that the switch remained down. However, there were several responses that described what the victim said during the procedure or how the shock generator was used, for example, it went up in 15-volt increments. To improve, candidates should know how all of the materials in the study worked to make everything appear to be genuine.

Question 4

(a) There were many strong responses that could clearly describe what participants had to do during their brain scans. Common descriptions included the number of scenes, how long they were projected for, and how the participant had to rate each scene. However, some responses focused on a different part of the procedure, for example, the recognition task. It is important for candidates to know all aspects of every Core Study on the syllabus to get a comprehensive overview of each one. It is also important to read the question carefully to ensure that a response is based on that part of the procedure.

(b) Many responses could identify an appropriate methodological weakness. These could be awarded the available mark by explicitly stating what the weakness is. Therefore, stating it lacked generalisability could gain the mark but simply stating generalisability by itself could not as the candidate had not identified what the weakness is. Other creditworthy responses focused on lack of ecological validity. It is important for candidates to be explicit in their responses to ensure that they are answering the question set and so the Examiner does none of the work. To improve on these types of questions, candidates need to be explicit in their responses as to why something is a strength/weakness.

Question 5

(a) A majority of responses could identify two characteristics of the sample. Common choices were them being candidates and having assumed high IQ scores. However, some responses gave features from another group in the Baron-Cohen *et al.* study, rather than Group 3. Some responses also stated males and females, but this was in the question so could not gain credit. To improve, candidates need to appreciate the difference between features of a sample and characteristics of a sample. Characteristics are the participant variables that are presented by the sample, whereas features are anything related to the sample, for example, characteristics, sample size, sampling technique etc.

(b) Stronger responses could clearly outline a potential problem and then put it into the context of the study by Baron-Cohen et al. Common responses included the eyes being static and that people could simply guess correctly. Many responses could outline a generic research methods based problem with the eyes test but then did not give an example based on the study itself, only gaining partial credit. To improve, candidates need to remember to explicitly link their responses to the study named in the question to be able to access all available marks.

Question 6

(a) Weaker responses tended to describe parts of the procedure from the study by Schachter and Singer. Stronger responses could clearly describe the two-factor theory of emotion sometimes with examples for elaboration. These answers demonstrated knowledge of the psychology being investigated in this study. To improve, candidates need to be able to describe key ideas and theories linked to each Core Study in isolation from the actual Core Study itself.

(b) For this type of question, responses must contain two parts. The first is an appropriate result. The second is an explicit link to the concept named in the question, in this case, two-factor theory of emotion. Some responses did present an appropriate result but did not then explicitly explain why the result could support the two-factor theory of emotion. Stronger responses demonstrated both parts of the question and could clearly link back to theory usually by explaining how the conditions/groups were set-up in the study.

Question 7

(a) Many responses could clearly outline what was meant by informed consent. However, there was a significant minority of responses that gave tautological answers, and these could not be credited. For example, a response may correctly give the ‘when the participant is given enough information’ part of the ethical guideline but finish with ‘and therefore the participant can give their consent’. Using the word to define the word cannot gain credit in an examination as it does not demonstrate that the candidate understands the full term. To improve, candidates need to be able to define key terminology without using the term itself.

(b) The majority of responses could identify one ethical guideline that was broken in the study by Laney et al. Stronger responses then used evidence directly from the study to outline why the guideline(s) had been broken. Popular choices were deception, psychological stress, and the lack of informed consent. Weaker responses tended to simply describe ethical guidelines without using evidence to show how it was broken. To improve, candidates need to read the question carefully as in this instance, the focus was on using evidence from the study.

Question 8

(a) Some responses could clearly outline the nature versus nurture debate used in psychology. However, a significant number of these, like with **Question 7(a)**, used the term to define the term. A common example was stating that nurture is based around nurturing a human being. This cannot gain credit as it does not demonstrate understanding. Stronger responses could give clear examples from studies other than Bandura et al. including amygdala activation (nature: Canli et al.) and learning through observation and the model/rival technique (nurture: Pepperberg). Some responses used Bandura et al. as the example and could not gain credit. To improve, like with **Question 7(a)**, candidates need to be able to define key terminology without using the term itself.

(b) Candidates need to appreciate the difference between a result and a conclusion. The former is factual and based on collected data. The latter is a generic comment based on the results reported in any core study. This question was asking about both of these. Weaker responses could give either a brief result or a brief conclusion from the study by Bandura et al. and nothing else. However, there were some stronger responses that could demonstrate knowledge of a conclusion from the study and then justify this with the data collected in the study. Popular choices included learning aggression via social learning and more likely from a same-sex model.

Question 9

The strongest responses evaluated the study by Piliavin et al. in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of qualitative data. Common choices included types of data collected, reliability, validity, generalisability, and ethics. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Piliavin et al. to explicitly support their point and tended to score Level 4 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses, all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by Piliavin et al. as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. In addition, some candidates followed a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). Therefore, some responses appeared to be prepared essays for Piliavin et al. without one of their points being about qualitative data. A response that does not have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did use qualitative data in an evaluative sense but did not fully explain <u>why</u> it could be a strength and/or a weakness. Stronger responses could identify the potential strength of gaining a more comprehensive explanation for helping/not helping and then give more than one example from the study making it 'in detail'. Several responses did not cover the named issue, only describing what qualitative data is or incorrectly evaluated the named issue by giving strengths/weaknesses of quantitative data. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue. Each strength and weakness should be of equal length with an explanation as to <u>why</u> 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. However, many candidates sitting this paper found demonstrating such understanding very challenging. In the absence of knowledge about fundamental concepts such as research methods, many candidates merely repeated the information in the question, for example saying that 'A questionnaire is a self-report method because the participant self-reports'. Similarly, in the absence of detailed knowledge about concepts, candidates often presented brief or irrelevant material that was connected to the question but did not answer it. For example, rather than outlining the concept of inter-rater reliability as required by the question, many candidates described how to test it or merely defined reliability.
- This paper expects candidates to apply their understanding to both familiar and unfamiliar contexts. Where some knowledge was evident, for example being able to identify a relevant study, candidates were often unable to link that knowledge to demonstrate why the study was relevant to the question.
- It is critical that candidates read and follow the instructions given in the question. These include reading and following the command words and applying concepts directly to novel scenarios. For example, when asked about problems of using a scale to measure memory, many responses focused on problems of studying memory rather than problems of using a scale as required by the question or made irrelevant comments about dreaming rather than memory. Similarly, in questions asking for definitions, candidates often repeated the question or gave an example in place of explaining what the term meant.

General comments

Candidates were often unable to access marks. Many responses suggested a lack of accurate knowledge and understanding or the skills needed to link responses to the scenarios, thus limiting performance as a whole.

Candidates across the ability range were able to demonstrate some knowledge of a range of aspects of research methods in this paper. Success was greater on some straightforward questions, such as graph plotting in **Question 4**, ethics in **Question 8(a)(i)** and identification of data type in **Question 9(b)**, than on more demanding questions, such as **Questions 9(c)(ii)** and **10**. However, there were also some more straightforward questions which candidates found challenging, such as **2**, **3**, **6** and **7(a)**.

Where questions require linking, such as to provide an example or to relate to a study, candidates were often able to earn partial marks for an initial identification of a relevant fact, such as laboratory and field experiments in **Question 6**, but were then unable to provide detail or link this information correctly to an example.

Although there were a few excellent answers to **Questions 10(a)** and **10(b)**, in general they were rarely well answered. In particular, understanding of 'semi-structured' was often confused, with responses suggesting that it referred to open and closed questions. Such responses therefore lacked the necessary relevant detail. To ensure that candidates are able to provide the necessary, relevant detail for each research method it is important that they have both the underpinning knowledge and have practised this high-level skill. It is most readily developed through practical work with designing and conducting small studies or through practice with novel scenarios.

Comments on specific questions

Section A

Question 1

(a) Part **(a)** was generally answered well, with most candidates able to identify that the data comes directly from the participant and all that data is written down. However, examples were more poorly done, with many candidates simply naming a study without giving any details of the questionnaire, thus not fulfilling the requirements of the question. Some identified studies such as Dement and Kleitman which did not use a questionnaire.

Responses which presented a circular argument, saying that 'A questionnaire is a self-report method because the participant self-reports', did not earn credit.

(b) This question part was not well answered. Many candidates suggested that strength of using questionnaires was that they could collect qualitative and quantitative data. This is an accurate fact but is descriptive rather than evaluative. Those candidates who attempted to add detail to this argument often suggested that interviews could not collect quantitative data, which is inaccurate.

Another common error was to suggest that using questionnaires was 'faster' or 'faster than interviews'. As a statement, this alone is insufficient, and needed justification. Such justification could, for example, have been given by saying that many participants can answer a questionnaire at the same time, so a large sample can be obtained and that this, in turn, would increase the generalisability of the results.

Question 2

This question was not well answered, with only a small number of responses offering clear null hypotheses. Many candidates said, 'no correlation' rather than 'no difference' and many showed little knowledge or understanding of the study.

Question 3

(a) Although here were some good answers, many candidates talked about the process of comparing rather than the agreement itself, i.e. outlined how to test for inter-rater reliability rather than outlining the concept.

Another common mistake was to define reliability rather than inter-rater reliability as required by the question.

(b) Most answers to **part (b)** were descriptions of aspects of the study unrelated to inter-rater reliability, so were not answering the question and could not be credited. There were, however, a small minority of very good answers describing the use of clearly specified categories as a way of ensuring inter-rater reliability. In addition, some candidates appropriately identified making observations at standardised intervals of five seconds as a factor improving reliability.

Question 4

Although this question was well answered by the majority of candidates. There were many indications of difficulties with graph drawing. Many candidates appeared not know what a bar graph is, for example drawing the bars touching, and axis labels often inaccurate or absent. Most candidates did attempt this question even when they left many other questions blank.

Question 5

This question was poorly answered with many responses focusing on problems of studying memory rather than problems of using a scale to measure memory. A small number of answers did appropriately discuss the lack of operationalisation or the subjective nature of the rating. In addition, there were many irrelevant answers here, with candidates referring to dreaming rather than memory.

Question 6

There were some good answers to this question although many described laboratory experiments as taking place in a laboratory and field experiments as taking place in a field (circular and incorrect respectively). In addition, candidates often incorrectly stated that laboratory experiment take place in a 'closed environment' or, conversely, (and also incorrectly) that field experiments are conducted in open environments or public places.

Many candidates confused field experiments with natural experiments. Another common error was to offer Milgram as an example of a laboratory experiment. Furthermore, many candidates simply named studies as examples rather than giving any details relating to their use as an instance of a laboratory experiment or as a field experiment.

Question 7

(a) The majority of candidates gave incorrect or incomplete answers to this question. Responses to this question part often scored one mark rather than two most as candidates simply identifying depth or detail. Many other candidates suggested a case study can collect 'more data', which could not be credited.

A very small minority of candidates explained that case studies can use multiple methods or can triangulate methods to improve validity, such as by using observations, interviews with the individual and questionnaires for relatives (or colleagues/friends).

(b) There were some good answers to **part (b)** although the link to the study was not always present and there were some irrelevant suggestions such as not actually meeting the woman or generic ethical issues. A small number of candidates simply stated terms such as 'lacks validity' or offered sampling/generalisability issues for both answers.

Question 8

(a) There were many good answers to **Question 8(a)**, but candidates sometimes confused question parts (between practical and ethical issues) and often offered consent for the first question implying that no consent would be required for a laboratory study. Practical issues for **part (ii)** were occasionally cost or time – non psychological issues – rather than relevant issues such as ecological validity. Answers were often written 'the wrong way round', so rather than explaining a benefit of using a laboratory the candidate gave a limitation of not using a laboratory. Such answers were credited, although answering the question explicitly is in the candidate's best interest, answering implicitly should be avoided. Many responses for the three **part (a)** questions were generic instead of being linked to this study.

(b) The answers to **part (b)** were generally good, with the majority of candidates able to explain that correlation is not causation and offer suggestions for other variables.

Question 9

(a) (i) Candidates demonstrated considerable confusion between participant/non-participant observation and covert/overt observation. This limited overall scores on **parts (a)(i)** and (**a)(ii)**.

(ii) This confusion resulted in many (incorrect) responses making reference to being covert or hidden, and the effect this would have on the chances of responding to demand characteristics, which is not directly relevant.

Even the better answers, offering relevant advantages, were often generic rather than applied to the context in **Question 9**.

(b) The majority of candidates identified 'quantitative' for the first mark in answer to **Question 9(b)** although a significant minority of responses stated that this was qualitative data. In general, this was where they went on to say that gender was not a number. A small minority of candidates simply said observational data.

(c) (i) Many candidates did not answer this question part well, often using the word 'situation' in their definition of a situational variable, and many simply said it was an uncontrolled variable. Other

responses did not answer the question, simply giving an example, such as 'the weather', rather than providing the meaning of the concept.

(ii) A small minority of candidates provided thoughtful answers, such as suggesting that the weather might be a situational variable causing helping to be difficult if people were carrying umbrellas or raincoats in the library making carrying books difficult. Some candidates described participant variables in **part (ii)** of this question and some described more general issues with the observation, neither of which were creditworthy. There was also a tendency for responses to relate to sampling, not to results, so no credit could be earned.

There was a further problem suggesting that the concept was poorly understood. Many candidates were suggesting participant variables or factors affecting the sample size rather than the influence of environmental variables on the behaviour or responses of participants.

Question 10

(a) This question part was not well answered although there were some good answers, in which candidates understood semi-structured interviews so were able to score very high marks. However, many responses suggested that candidates believe that 'semi-structured' refers to both open and closed questions, and some used the terms 'structured and unstructured questions'.

The problem of the research method was further illustrated by the number of candidates who unnecessarily, and often detrimentally, referred to – and described – irrelevant methods, including laboratory experiments, field experiments, natural experiments, questionnaires and case studies.

Candidates were often unable to progress beyond Level 1 and, even when candidates did give these details, they tended to forget to mention the types of questions (open or closed). This meant that they were unable to access the marks for detail about the 'style of questions' meaning that they could not access the Level 3 scores. Many candidates referred to experiments, independent and dependent variables, independent and repeated measures (because they asked more than one participant), observations and questionnaires, and random sampling when they meant opportunity. For example, common (incorrect) phrases were 'Random pedestrians', 'Random sampling of whatever people are walking by' and 'Pedestrians walking randomly on the street'. Overall, many candidates were not able to move out of level 1.

(b) **Part (b)** of this question was also poorly answered. It produced a significant number of answers focusing on sampling or ethics and many suggested the use of alternative (non self-report) methods. Many candidates needed to note the rubric saying: Do **not** refer to ethics or sampling in your answer.

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 the responses showed that the candidates were not well prepared for each of these skills and it would be beneficial to encourage them to prepare, 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 describing or evaluating, the answer must also be contextualised in a relevant way. Practice could help candidates to learn both how to extract relevant ideas, and how to make novel suggestions based on scenarios.

Question 10 in this paper requires candidates to produce an original design for a novel research question. This 'creative' process requires practice and it is, 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, 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 and, therefore, achieved the basic marks.

Some of the questions required a link, for example, to a study. These included **Questions 1(a)**, **1(b)**, **3(b)**, **5**, **7(e)**, **8(c)(i)**, **9(b)(i)** where candidate responses sometimes achieved partial marks for an initial identification of a relevant fact but failed to score the second mark because the answer was not linked. See **Questions 7** and **9** below for further detail. Candidates should always read the stem of the question and they should link any answer to the study the question is about. If the question is about a study being conducted, such as **Questions 7**, **8** and **9** the same applies: the answers to question parts should be related to the study being conducted.

There were some strong answers to **question 10**, although many responses brought in features of an experiment even though the question required the planning of a structured interview. This was not necessary, and, consequently, the response lacked the necessary relevant methodological features of a structured interview that were required to achieve higher credit. Many answers began with the selection of participants, often done in extensive detail. Whilst sampling is a relevant feature, it is only a minor feature and could be written about in less detail with more time allocated to the major features of the method.

Comments on specific questions

Section A

Question 1

(a) This question required the identification of the sampling technique. Many candidates achieved full marks by stating that it was volunteer sampling (or self-selecting sampling). Some candidates suggested incorrectly that it was opportunity sampling and a few, also incorrectly, claimed it was random sampling.

(b) This question required an explanation, and so required some detail in the answer. Answers scoring full marks explained that participants were taken from the subject pool because they had volunteered to participate in studies; that they were students from the University of Minnesota, or that they were introductory psychology students. Also creditworthy was that they gained two extra points in the final exam for participating. A few candidates wrote a general answer about volunteer sampling, and needed to focus their response as the question asked about the Schachter and Singer study. These answers could not be credited.

(c) Health checks were done on the participants to protect them from harm or because the checks avoided risks. Many candidates wrote answers like this and scored one mark. In order to achieve the second mark, the answer needed to be related to the Schachter and Singer study. Many candidates scored the second available mark by writing, for example, 'so they would not be harmed by the injection', or similarly 'so they would not be harmed by the effects of the adrenaline'.

Question 2

(a) Many candidates appeared to be looking for a more complex answer than the question required and were not all able to identify the sample and population from the information provided. Both question parts needed to be correct to score the 1 mark allocated to the question.

(b) Many candidates successfully scored full marks for this question part. Most candidates scored one mark for stating a reason, typically 'the sample was not representative', but again, the answer had to be related to the study to score full marks. Candidates who scored full marks wrote answers like 'the sample was not representative because the 72 children in the study were from the same Stanford nursery school probably with intelligent, wealthy parents, which is not typical of most other children.'

Question 3

(a) Subjectivity is the effect of a personal viewpoint/opinion/bias on data collection or interpretation. Candidates stating this, or words to the same effect, scored the one allocated mark.

(b) Answers to this question illustrated the need for candidates to relate their answers to the study in question. The words 'in (or for) this study' is the trigger to do this. If the 'in this study' component is absent, then full marks cannot be awarded. For this question, most answers suggested that qualitative data could be gathered by using an interview or questionnaire, and one mark was awarded. Some answers went no further; other answers suggested using an open question but then gave an example of a closed question (which gives quantitative data). Answers scoring the further mark suggested: 'ask an open question such as *describe* how you feel on your birthday'.

Question 4

(a) Some candidates did not score marks for two reasons: they did not know what a standard deviation was and/or they did not relate standard deviation to the results of the Baron-Cohen et al. study. Many candidates merely *described* the words and numbers in the table (e.g. the AS/HFA group scored 6.0 which is higher the IQ group score of 2.9) rather than *explaining* what the results meant. An answer scoring full marks might be, for example: 'standard deviation is about the spread of scores from the mean, so the AS/HFA group score of 6.0 shows they were more spread out in their AQ scores than the IQ matched group'.

Question 5

(a) This question required three components to score the available three marks: an explanation of the term 'operationalisation', how this applied to the IV and how this applied to the DV of the Piliavin et al. study. Many candidates provided very strong answers and scored full marks. The term operationalisation means clearly defining a variable so it can be accurately manipulated, measured or replicated. In this instance the IV was drunk or ill, with drunk being operationalised as carrying a bottle in a bag, and ill being operationalised by the carrying of a cane. The DV wasn't just 'helping', it was operationalised as spontaneous helping, which is helping before the model.

Question 6

Most candidates were able to describe the terms reliability and validity successfully. A number of candidates opted to give further definitions of the different types of reliability and validity, and this was generally a rewarding strategy. For example, some described test-retest reliability and inter-rater reliability. For validity some described concurrent validity, and ecological validity was often included. An alternative strategy was to provide examples from studies, and relevant studies such as that by Bandura et al. were used. The most successful strategy, which often earned full marks, was to give definitions of the terms, describe different types and then give examples of the types.

Question 7

Candidates should always be mindful that the introduction to the question, the stem, applies to all question parts which follow.

(a) Nearly all candidates answered this question part correctly when stating that the IV was the 'lighting level', 'bright lighting and dim lighting' and even 'lighting' was sufficient for one mark. A few candidates stated the DV which was incorrect.

(b) Nearly all candidates were awarded one mark for their DV answer. Most typically 'attention' and 'score on the attention (or listening) task' were creditworthy. A few candidates repeated the stem 'the effects of lighting on attention' and a few candidates provided the IV.

(c) All three question parts, **(c)(i)**, **(ii)** and **(iii)** are linked. Candidates suggesting a relevant variable for **(i)** nearly always went on to suggest how this could be controlled **(ii),** and why it should be controlled **(iii)**. The suggested variable could be of any type, such as a situational variable or a participant variable. An answer scoring full marks might be **(i)** a situational variable, such as 'background light would need to be controlled. **(ii)** This might be done by all participants being tested in the same darkened room and this **(iii)** would ensure that the IV of bright and dim lighting was tested and not an extraneous variable of differing amounts of background light.

(d) Many candidates scored full marks when suggesting that a listening task would be appropriate because the lighting was either bright or dim, and the light might be an extraneous variable if a test using light was the DV. Some candidates suggested that people have to concentrate more on a listening task. Those answers scoring full marks referred to how the validity of the test might be affected by their suggestion.

(e) Answers to this question could be categorised into three types: **(i)** those who could not suggest a weakness; **(ii)** those who suggested that being in a laboratory would reduce ecological validity (the task would be artificial) and these answers scored one mark; **(iii)** those who gave a reason, such as that in **(ii)** above, but then went on to relate the reason for Zayn's study, by writing, for example 'a listening task in bright or dim light is not something a person would do in everyday life.' Answers like this scored the full two marks.

Question 8

(a) Candidates had to draw a histogram, and only a few candidates managed to score the three available marks. Marks were awarded for correctly labelling the *x*-axis (the categories of temperature) and the *y*-axis (the frequency of each category). This was done successfully by many candidates. Additionally, the data had to be plotted. A histogram requires the bars to be the same width and have no gaps between them. Whilst a few candidates did this, many included gaps, plotted the data incorrectly and some attempted to plot the data in the form of a correlation. All these answers could not be awarded the additional mark.

(b) Acceptable answers for this question were that 'hot weather makes people more angry' or 'people speed more when the weather is warmer'. What was not acceptable was an answer that referred to a correlation, such as 'there is a positive correlation between speeding offences and temperature'. Data in the form of a histogram can never be 'converted' into correlational data, conclusions about a histogram cannot become correlational.

(c) (i) Some candidates did not score marks for two main reasons: the variable suggested was not a situational one as the question required; or the suggested variable was temperature which was not allowed because all of question 8 is about temperature (i.e. the data in the stem and histogram). However, the effect of temperature on some other variable was acceptable. For example, a very hot temperature might reduce amount of traffic on the roads (so the situational variable is traffic density). Many candidates also failed to answer the second part of the question, which was to consider 'how this could have affected Gwyn's results'. A full mark answer would be, for example: 'there might be less traffic on the roads when it is hot (one mark) which means that people might speed more' (two marks).

(ii) For candidates correctly answering part **(c)(i)** the answer to this question part was an obvious consequence and nearly all candidates scored the one available mark. For example, stating 'there might be less traffic on the roads when it is hot which means that people might speed more' led to the answer 'to count the number of cars and only collect data on days of similar traffic density'.

Question 9

As mentioned in relation to **Question 7**, it is essential that the stem of the question is read before beginning to answer the question parts.

(a) (i) The main weakness was that many candidates suggested that the elderly people would open-up to Hazel and answer any of her questions. However, the stem states that the study is an observation and so there would be no questions, leading to such answers scoring 0 marks. Correct answers would be that the participants would behave normally because she is familiar to them rather than how they would behave if a stranger were observing them.

(b) (i) Most candidates were awarded one mark for referring to ethical problems such as the risk of harm, invasion of privacy or reduced right to withdraw. What was often absent is how the ethical problem related to Hazel and her study. Candidates scoring full marks included answers such as (for the right to withdraw): 'the elderly people might feel that they are unable to withdraw because they are expressing emotions because Hazel knows them'.

(ii) One mark was awarded for a general answer, such as 'make the right to withdraw clear' and one further mark was awarded for suggesting how Hazel could apply this in her study. One way would be to state: 'Hazel could put up a sign explaining to residents that they could go to their rooms whenever they wished'. For potential psychological harm, debriefing could be done (one mark) and Hazel could reassure a participant if she sees them becoming distressed (+ one mark).

Question 10

(a) There are many methods in psychology and that only an experiment has an IV and DV. Other methods such as observations, interviews and questionnaires are not experiments and so 'experimental terminology' does not apply. This question required candidates to conduct a study using a structured interview and as this is not an experiment, any mention of experimental terminology was irrelevant and scored no marks. Many candidates included the major features of a structured interview in their suggested studies, which included the content of the questions (topics such as serious/playful, optimistic/pessimistic); how the structured interview would be conducted (same questions to all and in the same order); the style of the questions (open, closed or both). In addition, marks were awarded for minor features such as the location of the study, the participants, how answers to questions would be scored. Answers addressing all or most of these features often scored very high marks; some responses at the bottom end of the mark range addressed only one or two and often in a vague way. A small number of candidates designed an experiment rather than conducting a structured interview, which could not be credited.

(b) Although there were many maximum mark answers, there were others that did not score full marks. There were several reasons for this. **(i)** Despite the question stating 'do not refer to ethics or sampling', a number of answers incorrectly focused on one of these aspects. **(ii)** Some answers

were very brief, often no more than a single sentence, when more detail could have scored more marks. **(iii)** Confusion concerning a structured interview. Many candidates incorrectly referred to a structured interview only gathering quantitative data (and an unstructured interview gathers only qualitative data). A structured interview simply means that all participants get the same questions in the same order, and the questions asked can be open or closed.

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. Such questions require different skills, some of which presented difficulties for some candidates. It is essential that candidates are prepared for each of these skills, especially understanding terms and the accurate application of this knowledge.
- Applying knowledge and understanding to novel scenarios is important to succeed on this paper. This could have helped candidates in two ways:
  - Candidates needed to be able to apply research methods terms and concepts to scenarios presented in questions. These can include, for example, planning, criticising or developing designs or analysing data.
  - Candidates must take note of questions which indicate the need for examples or for a link. When a question says 'in this study', makes direct reference to the scenario, or states 'give an example', responses must go beyond simply describing or evaluating. Candidates therefore need to be prepared for questions using such formats and practice can help them to learn both how to extract relevant ideas and how to make novel suggestions.
- It is critical that candidates read and follow the instructions given in the question. These include reading and following the command words, applying concepts directly to novel scenarios when these are given and reading the whole question.

General comments

Candidates were able to access marks across the whole paper. However, not all were able to accurately and consistently demonstrate knowledge and understanding, or access the additional marks for linking their response to the scenarios, thus limiting performance as a whole.

Candidates across the ability range were able to demonstrate some knowledge of a range of aspects of research methods in this paper. Success was greater on some straightforward questions, such as **Questions 1(a)**, **3**, **7(a)**, **(b)** and **(c)**, **8(b)(i)** and **(ii)** and **9(a)**, than on more demanding ones, such as **Questions 7(d)** and **10**. However, there were also some more straightforward questions which candidates found challenging, such as **8(a)**, **9(b)** and **9(c)**.

Where questions require linking, such as to provide an example or to relate to a study, candidates were often able to earn partial marks for an initial identification of a relevant fact, such as types of observations in **Question 6**, but were then unable to link this information correctly to an example.

Question 10 was sometimes well answered although responses were at least partly irrelevant and often gave inappropriate or incomplete examples of open and closed questions – an essential part of a self-report such as a questionnaire. Such responses therefore lacked the necessary relevant detail. To ensure that candidates are able to provide the necessary, relevant detail for each research method, it is important that candidates have practised this high-level skill. It is most readily developed through practical work with designing and conducting small studies or through practice with novel scenarios.

Comments on specific questions

Section A

Question 1

(a) This question part was generally well answered, with many candidates achieving at least one mark by stating tiredness or boredom. A common error here was to describe fatigue effects by using the word fatigued – it is important that candidates do not write circular statements, i.e. candidates should not repeat the key word in their definitions.

(b) A range of acceptable responses were given here. However, several candidates simply stated that Alex made more errors, which was incorrect.

Question 2

Most candidates earned the mark here. However, many experienced a similar problem to **Question 1(a)**, i.e. one of repetition. On this occasion, some candidates simply repeated that an opportunity sample lacks representativeness without explaining why.

Question 3

Many candidates achieved full marks on this question, with the most common responses referring to the juice reward and linking clearly to a lack of punishment/deprivation.

Some responses referred merely to having space.

Question 4

(a) Many candidates were able to access the full 4 marks. However, a significant minority of candidates did not address the end of the question stating it had to be data other than from the fMRI. More commonly, candidates could identify the two quantitative measures but did not give the detail necessary for the second mark or, alternatively, gave the detail for remembered / forgotten / familiar but did not state that the quantitative measure was the number of slides recalled / remembered.

(b) (i) Many candidates found the precise nature of application of knowledge required by this question difficult, with responses focusing on the study rather than the measure. However, there were a few excellent responses giving good detail from the original study with regards to the correlation between valence and emotional intensity ratings.

(ii) Candidates were better able to access the marks on this question part than on **(b)(i)**, with a common response being 'individual differences' or 'participant variables' and the interpretation of the emotional ratings.

Question 5

A common response to this question was to give a generic statement of 'to avoid demand characteristics'. However, the mark required that the answer was specifically applied to Andrade's study, rather than merely being the identification of a term.

Question 6

There were some strong answers to this question but also significant errors. Many candidates appeared to be confused between covert and overt observations. In addition, there were many responses which referred to participant/non-participant observation rather than overt/covert. Some responses suggested that covert observations were the same as participant observations and that overt observations were necessarily non-participant.

Examples often just stated a reference without giving specific detail that illustrated how it was relevant to the question (e.g. 'Piliavin' was not enough, the response needed to say, for example, that the observers were disguised as passengers).

Section B

Question 7

(a) Most candidates performed well on this question. Nevertheless, some candidates only gained one mark as their second point was just a reversal of the first (e.g. count how many shapes sorted correctly (first point), count how many shapes were sorted incorrectly (second point).

In addition, a number of candidates identified different methods (e.g. observation, self report), i.e. did not answer the question.

(b) Many candidates were able to indicate that they recognised the need to change the way consent was gained from children, although some simply presented their answer as though it was intended for an adult, thus could not earn credit. Some strong responses included Dhia demonstrating sorting to the children so that they could decide and telling them that she would 'score the game'.

(c) This question part was mostly well answered with a range of ideas that were appropriate to children.

(d) This question part was not as well answered as the rest of **Question 7**. This was in part because many candidates were using the terms validity, reliability and generalisability interchangeably and simply stated 'this will affect the validity and reliability of Dhia's results' without explaining the effect.

Candidates need to show clear understanding of the different terms, and to be able to apply them to novel situations.

Question 8

(a) This question part did not receive many strong answers. A common error was for candidates to respond that interviews will allow for qualitative data to be gathered. However, questionnaires can also easily gather qualitative data so this was not a reason.

(b) (i) This question part was generally well answered with a range of appropriate reasons being offered, for example that people with back pain may spend more time in bed and so be exposed to less sunlight.

(ii) This question part was also generally well answered with a range of appropriate reasons. For example, selecting people with hand pain, because they would be more mobile than patients with back pain or using athletes who have sustained injuries as they would normally be outdoors in the sun. In some cases, candidates merely stated a different sampling technique, which was not answering the question so could not earn credit.

(c) Few candidates marked units on the sunlight axis and there were many examples of positive correlations and of bar charts.

Question 9

(a) This question part was generally well answered, with 2 marks typically being awarded for two suggestions rather than one suggestion with detail.

(b) This question part did not receive many strong responses. There were many references to correlations not equalling causation which were not relevant to the question. Other candidates identified that there may have been participant variables in terms of memory but did not clearly contextualise this to potential differences between the groups of doodlers and non-doodlers.

(c) This question part did not receive many strong responses. A significant number of candidates incorrectly referred to a debrief as prior to a study to gain informed consent, hence they could not earn credit. Some candidates correctly identified returning participants to their original state but did not link this idea to Mitesh's study (e.g. feeling bad that they hadn't recalled many noises).

Section C

Question 10

(a) Many candidates made their responses to this question much more complex than was required, for example describing complex laboratory experiments/field experiments/covert observations/stooges at parties. Many responses were limited to a Level 1 score.

Candidates often appeared unaware that they should include response options for closed questions (e.g. Yes/No). Many responses, however, did include reference to qualitative and quantitative data, so were able to access Level 2. A minority of candidates gave closed questions as examples of open questions. Nevertheless, there were some excellent responses.

(b) Strong responses here are based on the details of the study they have described in **part 10(a)**. However, candidates need to ensure that they are taking notice of guidance in the question (e.g. not referring to ethics or sampling) and must clearly link their problems/solutions to their procedure. Only a few candidates did this.

PSYCHOLOGY

Paper 9990/31
Specialist Options: Theory

Key messages

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

It is important that candidates are made aware of the terminology, concepts and theories identified in the syllabus, as well as key terms used in named theories and studies, as some were unable to identify and/or define the terms given in these types of questions. Creating a glossary of key terms, revision of terminology using flash cards and class quizzes on terminology could prove useful. Where the response gave an example to help define the term or theory, it often achieved full marks. These questions are worth 2 marks and a brief response is appropriate.

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

These questions could ask the candidate to describe a theory, study, treatment or technique 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 the findings, or a summary of the key features of the study. This question is worth 4 marks and the candidates should write a more extended answer. It would be helpful for candidates to create a revision flashcard or mind map of each bullet point in the syllabus. The flashcard should be given the title used in the syllabus, for example, 'treatment and management of schizophrenia and delusion disorder: cognitive-behavioural therapy: Sensky', to help the candidate to identify which part of the syllabus the question is referring to. If the question asks for a part of the study, the response should only describe this part. For example, if the question asks for the procedure, then the response should not describe the findings of the study.

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

These questions could require the candidate to explain up to two strengths/advantages or weaknesses/ disadvantages 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 such as the effectiveness of controls used in a study. 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 named in the question. To improve, responses should give specific examples to support their point. As mentioned for the odd question **part (b)**, the candidate could make a flashcard/revision notes and include strengths and weaknesses of the theory, study or technique to help candidates prepare for these questions.

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

This question comes from one of the bullet points in the syllabus. Candidates could describe the three or four studies, theories or techniques identified in the syllabus under the appropriate bullet point. It is possible for the responses to achieve full marks by describing at least two of the studies, theories or techniques and this would need to be a very detailed description. Full marks can also be achieved by a response that describes three of the bullet points in detail (in less depth than if the response described two of the studies, theories or techniques) with excellent understanding and good use of terminology throughout. It is also important that the descriptions are linked to the topic area named in the syllabus. It could be useful for candidates to create revision notes with the title of each bullet point as the header in their notes.

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

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

Some of the responses covered other issues rather than the one named in the question. Quite a few of the answers were structured by study/theory/technique rather than by the issue, which often led the response to be quite superficial and repetitive. A few of the responses did analysis. Candidates should be aware this question is worth 10 marks and attempt to include an appropriate amount of information.

General comments

The marks achieved by candidates for this session of the 9990 syllabus achieved across the full range of the mark bands.. Some candidates were well prepared for the exam and showed good knowledge, understanding and evaluation throughout their responses. A number of candidates were not as well prepared and showed limited knowledge and understanding with brief and/or superficial responses. These candidates often showed limited evaluation skills.

Time management for this paper was good for the majority of candidates and most attempted all questions that were required. Some candidates did not respond to one or more of the questions asked in the option area. A very small number of the candidates attempted to respond to more than two topic areas but often did not attempt all 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) Many of the responses addressed the question and achieved one or two marks by outlining a cognitive explanation of schizophrenia and delusional disorder. Common responses included outlining faulty reasoning, difficulties with self-monitoring, and a failure to recognise hallucinations as that of the sufferer's own voice. Some responses outlined a behavioural or biochemical explanation of schizophrenia which was not creditworthy.

(b) A significant number of responses were able to describe the procedure in the study by Sensky et al. of cognitive-behavioural therapy (CBT) for schizophrenia. Strong responses referred to the participants, the random allocation to treatment, details of CBT and befriending treatments, as well as some details of how the patients were assessed. Weaker responses gave an outline of CBT with no other details of the study provided. Some responses described the incorrect study such as the study on token economies by Paul and Lentz. These types of responses were not creditworthy.

(c) The responses to this question covered the full range of the mark scheme. Stronger responses outlined both a similarity and a difference between CBT and a biochemical treatment for schizophrenia/delusional disorder. Most common good similarities looked at intervention from therapist or evidence for success, good differences included the fact that the treatments were

based upon different approaches and prevalence of side effects in drug treatments but not CBT. Weaker responses often identified the similarity or difference without providing any examples. Some responses compared CBT to ECT which is not a biochemical treatment. These types of responses were not creditworthy.

Question 2

(a) Responses varied considerably for this question and covered the full range of the marks available, although most achieved in the lower mark bands due to lack of detail in the response and/or errors in the description.

Stronger responses gave good definitions of bipolar and unipolar disorder as well as detailed outline of the BDI. Weaker responses described bipolar disorder but stated that the change from depression to mania could happen very quickly and last a very brief period of time which is incorrect. There were a number of descriptions of explanations and/or treatments and also a significant proportion of candidates who described a number of disorders from other parts of the syllabus, notably schizophrenia, phobias and OCD. These types of responses were not creditworthy.

(b) Many of the responses achieved in the Level 1 or Level 2 mark band. A few candidates provided clear analysis and details of the characteristics and measures to back up their evaluative points that enabled their responses to achieve Level 3 and above. There was a tendency for responses to focus on many issues per characteristic/measure rather than applying the issue to the different characteristic/measure.

A few responses did effectively discuss the named issue of validity with some clear analysis. Many of the responses did not address the named issue of validity. Common evaluation issues included self-reports, reliability and application to everyday life. Some responses just evaluated the BDI with no attempt made to evaluate the characteristics and the issues surrounding diagnosis. Some responses evaluated explanations and/or treatments in **part (b)** which was not creditworthy.

Psychology and Consumer Behaviour

Question 3

(a) There were a few strong responses to this question with an identification of two stages of the consumer decision model other than 'recognition of need'. Popular stages included identifying 'informational search' and 'evaluation of alternatives'. Weaker responses identified one stage correctly. Many candidates did not attempt this question or gave incorrect stages.

(b) There were some very clear and detailed responses describing two different ways to close a sale. Common responses included reciprocity, liking, scarcity and body language. Responses that gave examples to explain how the way to close a sale would lead to a purchase often achieved full marks. There were some candidates that did not attempt this question.

(c) There were some strong responses to this question with the vast majority being able to give both an advantage and a disadvantage of one of the ways to close a sale outlined in **part (b)**. Common advantages included explaining why this way of closing the sale would lead to a purchase and increased income for the company, repeated sales in the future through word of mouth and increased reputation of the company/product. Common disadvantages included explaining why the consumer might not purchase the product due to feeling pressured and lowered reputation of the company. Strong responses identified the advantage/disadvantage and gave clear examples to explain their point. Weaker responses did not explain with any clarity about why it was a strength or weakness. For example, many simply outlined that the company would have increased sales rather than looking at the point of view of the consumer and their satisfaction. These types of responses achieved in the Level 1 mark band.

Question 4

(a) Many responses achieved Level 1 or Level 2 for this question. A few were able to achieve in Level 3 or Level 4 by giving fairly detailed descriptions of the models of the effects of ambience including pleasure-arousal and cognition-emotion, the Kutlu et al. study on lighting and colour in retail stores and the Chebat and Michon study on the effects of odour on shopper arousal and emotion. Weaker

responses often briefly described the studies by Kutlu et al. and the study by Chebat and Michon but often with few details. Some responses just outlined the findings/conclusion of these studies with no other details given. These types of responses often achieved in the Level 1 mark band.

(b) Those responses that achieved in the higher mark bands for **part (a)**, tended to produce good answers to this question, with some understanding of how control of variables applied to the studies with some examples given. Some attempted analysis, for example pointing out that samples in both the studies lacked generalisability with some examples given to back up the point made. Weaker responses did not attempt the named issue of control of variables. This appeared to be because as the study was not described in **part (a)**, the candidate was not aware of the controls used in the studies. Ecological validity, ethics and self-reports were other commonly chosen issues. Weak responses often named an issue and then stated that a study either supported the issue (had good ethics) or did not support the issue (had poor ethics) without any explanation given for the point made. These types of responses achieved in the Level 1 mark band.

Psychology and Health

Question 5

(a) Many responses achieved 1 mark by giving examples of the features/components of the Yale model of communication such as the communicator, target, attention, etc. However, most responses did not outline the model in relation to health promotion to achieve full marks. A few responses did this by giving an example related to health promotion for one of the features, such as the communicator being a well-respected doctor/expert in the field of medicine, and achieved full marks. A number of responses did not know the Yale model of communication and either did not answer the question or wrote a general response about health. These types of responses were not creditworthy.

(b) There were a few strong responses to this question with some giving a detailed description of the Lau et al. study on health change in adolescents. Common responses included the aim, that the study was longitudinal, that adolescents and their parents answered questionnaires and some of the results of the study. Weaker responses often outlined that the study used a questionnaire and gave a result. There were many incorrect responses with candidates outlining a study done on children rather than adolescents. These types of responses were not creditworthy.

(c) Where there was a creditworthy response in **part (b)**, the candidates were often able to outline both a strength and a weakness of the Lau et al. study. Common strengths included generalisability of the sample size, practical application and being able to monitor change in health belief/change over time. The most common weakness was generalisability due to the participants being from one university. Many responses achieved Level 1 or Level 2 due to not providing any clear example from the Lau et al. study to back up their strength and/or weakness. Responses that were not creditworthy typically evaluated the incorrect study.

Question 6

(a) The responses to this question covered the full range of the mark scheme. Stronger responses had a good understanding of the measures of pain and focused on 3 or 4 measures rather than attempting to describe all of the measures detailed on the syllabus. There were some very detailed descriptions of the McGill Pain questionnaire, UAB pain behavioural scale and the Wong-Baker scale. Weaker responses attempted to describe a large number of measures of pain but very superficially (and/or incorrectly). These types of responses typically achieved in the Level 1 mark band. A number of responses described theories of pain rather than measures of pain which was not creditworthy.

(b) There were some strong responses to this question. These were often able to evaluate the named issue of psychometrics and give the strengths and weaknesses of this type of measure, as well as examples from the measures of pain described in **part (a)**. Other common evaluation issues for this question included the reliability of the measurements, self-reports (although this sometimes just repeated the issues of reliability and/or validity), quantitative data and objectivity versus subjectivity. Weaker responses seemed to be when a large number of issues were covered but all were done in a very superficial manner, simply stating that issues did or did not apply to the measurements but not why. These types of responses often achieved Level 1.

Psychology and Organisations

Question 7

(a) There were some strong responses to this question with some achieving full marks. The most common practices identified from Kouzes and Posner's Leadership Practices Inventory (LPI) were 'encourage the heart' and 'model the way' which was either identified or described. Weaker responses identified one of the practices (typically 'model the way'). Some responses gave an outline of a theory of leadership which was not creditworthy.

(b) There were several good, detailed responses to this question. There were some strong answers describing the styles of leader behaviour proposed by Muczyk and Reimann including the dimensions of directive/permissive and autocrat/democrat. Other good responses described specific types e.g. directive autocrat. Some responses confused the styles of leader behaviour with situational leadership. These types of responses were not creditworthy.

(c) There were a few good responses to this question with both a similarity and a difference given between Muczyk and Reimann's styles of leader behaviour and Hersey and Blanchard's situational leadership. Stronger responses outlined the similarity that both depend on the situation or require the leader to be flexible/adaptable, and some did give examples from both theories to back up their similarity. A common difference was that the styles of leader behaviour sees a distinction between making decisions and carrying them through whereas situational leadership does not. If the candidate outlined situational leadership in **part (b)** of their response, they were often unable to outline either a similarity or a difference.

Question 8

(a) There were some good, detailed responses to this question with clear details given of group development and roles in organisations, including Tuckman's theory of group development, Belbin's theory of team roles and Belbin's team inventory. Tuckman and Belbin's theories tended to be the most detailed of these descriptions. Weaker responses tended to be brief or gave anecdotal descriptions of group development and roles. A number of responses referred to other sections of the syllabus, such as attitudes to work and theories of motivation. These types of responses were not creditworthy.

(b) There were a few good responses to this question where it was structured by evaluation issue and began with the named issue of practical applications. Some of these responses gave clear examples from **part (a)** of their response to back up their evaluative points and attempted some analysis. In addition to practical applications, popular evaluation issues were individual and situational, and reductionism versus holism. Weaker responses did attempt to evaluate the named issue but did not explain why the research had a practical application and did no analysis. These types of responses tended to state that a theory or study either was or was not applicable to companies with no explanation of this given. Those responses that attempted individual and situational explanations were often weak with confusion over the definition of this issue.

PSYCHOLOGY

Paper 9990/32
Specialist Options: Theory

Key messages

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

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

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

These questions could ask the candidate to describe a theory, study, treatment or technique such as a 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 the findings, or a summary of the key features of the study. This question is worth 4 marks and the candidates should write a more extended answer. It would be helpful for candidates to create a revision flashcard or mind map of each bullet point in the syllabus. The flashcard should be given the title used in the syllabus, for example, 'characteristics of anxiety disorders: measures: Generalised Anxiety Disorder assessment (GAD-7)', to help the candidate to identify which part of the syllabus the question is referring to. If the question asks for a part of the study, the response should only describe this part. For example, if the question asks for the findings, then the response should not describe the procedure of the study.

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 such as the reliability of a study. 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 named in the question. To improve, responses should give specific examples to support their point. As mentioned for the odd question **part (b)**, the candidate could make a flashcard/revision notes and include strengths and weaknesses of the theory, study or technique to help candidates prepare for these questions.

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

This question comes from one of the bullet points in the syllabus. Candidates could describe the three or four studies, theories, 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, disorders or techniques and this would need to be a very detailed description. Full marks can also be achieved by a response that describes three of the bullet points in detail (in less depth than if the response described two of the studies, theories or techniques) with excellent understanding and good use of terminology throughout. It is also important that the descriptions are linked to the topic area named in the syllabus. It could be useful for candidates to create revision notes with the title of each bullet point as the header in their notes.

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

This question asks the candidate to evaluate the theories, studies, disorders and/or techniques described in **part (a)** of the question. The response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. However, most responses that evaluated using two issues in this exam, achieved marks in the lower bands due to the response being superficial and often with little analysis. Some responses that considered three issues tended to achieve higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the theories, studies, disorders and techniques described in the **part (a)** of the answer. The candidate must also provide some form of analysis. This could be done by discussing the strengths and weaknesses of the issue being considered, presenting a counterargument 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. To achieve the requirements of the Level 3 and 4 band descriptors, ideally the response should be structured by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure the answer covers this requirement of the question.

Some of the responses covered other issues rather than the one named in the question. Quite a few of the answers were structured by study/theory rather than by the issue, which often led the response to be quite superficial and repetitive. A number of the responses did analysis. Candidates should be aware this question is worth 10 marks and attempt to include an appropriate amount of information.

General comments

The marks achieved by candidates for this session of the 9990 syllabus achieved across the full range of the mark bands. Many candidates were well prepared for the exam and showed good knowledge, understanding and evaluation throughout their responses. Some candidates were not as well prepared and showed limited knowledge and understanding with brief and/or superficial responses. These candidates often showed limited evaluation skills.

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

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

Comments on specific questions

Psychology and Abnormality

Question 1

(a) There were some strong responses to this question with an outline given of classical conditioning and how this leads to a phobia. Some responses did include appropriate terminology such as outlining how the neutral stimulus is paired with an unconditioned stimulus which will eventually lead to the conditioned response of fear to the conditioned stimulus. A number of responses gave the example of Little Albert and how he learned to be frightened of a white rat and these often achieved full marks. Weaker responses tended to identify classical conditioning but were then unable to explain the conditioning process or explained it incorrectly. Some responses gave a definition of phobias with examples which was not creditworthy.

(b) A significant number of responses were able to describe the Generalised Anxiety Disorder assessment (GAD-7). Common features included that it is a screening tool, it measures severity of anxiety, consists of 7 items, and has a 0–3 scale. A few responses were able to include an example of the statements used in the assessment. Weaker responses included fewer details of the assessment. Some responses incorrectly identified that this assessment measures the severity of a phobia rather than anxiety disorder which was not creditworthy.

(c) Responses to this question covered the full range of the mark scheme. Stronger responses identified a strength and a weakness and explained this in terms of the GAD-7 with a clear example. Common strengths included quick and easy (often linked to it being a screening tool so that once assessed the patient can be referred to a specialist for further assessment) and strengths of quantitative data collected such as making a comparison before and after treatment. Common weaknesses included social desirability, lack of quantitative data and that the GAD-7 is just a screening tool and cannot be used for diagnosis of generalised anxiety disorder. The weakness of social desirability was often not very well explained. Although this weakness was creditworthy, it should be noted that a self-report is the best measure to find out how a person is feeling and it is unclear why a patient would not tell the truth in this assessment. Weaker responses had a lack of depth in their explanations and gave no example from the GAD-7 to back up their point.

Question 2

(a) Responses varied considerably for this question and covered the full range of the marks available. Some of the candidates were very well prepared for this topic, whereas others showed very limited knowledge. There were some detailed responses which were accurate and coherent with a good use of psychological terminology. The strongest responses focused on describing the symptoms of schizophrenia and psychotic disorders. Many were able to identify which symptoms were positive and negative and gave good definitions of each symptom and type of delusions. Many gave clear descriptions of the Freeman study on symptom assessment using virtual reality. Weaker responses sometimes outlined the symptoms of schizophrenia without identifying which are positive and negative or giving any description of the symptoms. Some responses did not describe the Freeman study or gave a brief/inaccurate description. These types of responses often achieved in the Level 1 mark band. A number of responses outlined explanations and/or treatments for schizophrenia and delusional disorders and these types of responses were not creditworthy.

(b) Many of the responses achieved in the Level 1 or Level 2 mark band, with a few providing clear analysis and details of the characteristics and the Freeman study to back up their evaluative points that enabled these types of responses to achieve Level 3 and above. There was a tendency for responses to focus on many issues per characteristic/measure rather than applying the issue to the different characteristic/measure. A few responses effectively discussed the named issue of validity with some clear analysis. Some of the responses did not address the named issue of validity. Common evaluation issues included reliability and application to everyday life. Some responses just evaluated the Freeman study with no attempt made to evaluate the characteristics and the issues surrounding diagnosis. Some responses evaluated explanations and/or treatments in **part (b)**, which was not creditworthy.

Psychology and Consumer Behaviour

Question 3

(a) There were some strong responses to this question with a clear outline of the AIDA model of advertising. Some were able to identify attention, interest, desire and action and could achieve 1 mark. Some responses gave an example and linked the example to advertising to achieve the 2nd mark. Some responses gave a general outline of what a company should do to advertise a product and these types of responses were not creditworthy.

(b) There were some clear and somewhat detailed responses describing the study by Kohli et al. on effective slogans. Some showed good knowledge of the 7 recommendations alongside some understanding of why the recommendations were put forward. Weaker responses often outlined one or two of the recommendations. There was a lot of misunderstanding that this study was not an experiment but a review article of other studies. Some responses outlined a study on slogans that was not from the Kohli et al. study and was therefore not creditworthy.

(c) Where there was a creditworthy response in **part (b)**, the candidates were often able to outline at least one strength of the Kohli et al. study. Common strengths included practical applications and validity. Many responses achieved Level 1 or Level 2 due to not providing any clear example from the Kohli et al. study to back up their strength. Responses that were not creditworthy typically evaluated the incorrect study or used an inappropriate evaluation issue such as reliability or ecological validity, without any reference to one of the studies outlined in the Kohli et al. study. These types of responses were not creditworthy.

Question 4

(a) Many responses achieved Level 2 or 3 for this question. Good descriptions were often seen in responses for the Porublev et al. study on gift wrapping. There were also some good descriptions of product colour and associative learning by Grossman and Wisenblit. The descriptions of the Atalay et al. study on attention and shelf position were often brief or gave incorrect details such as stating that the second study took place in a supermarket. Weaker responses gave very brief or inaccurate details of the research which achieved in Level 1. Some responses just described different colours of products and how consumers might respond to them which was not creditworthy.

(b) Those responses that achieved in the higher mark bands for **part (a)** tended to produce strong answers to this question, with some understanding of how self-reports applies to the studies with examples given. Some attempted analysis, for example pointing out the strengths and weaknesses of self-reports with examples given from the research outlined in **part (a)**. Generalisability, ecological validity and applications to everyday life were other commonly chosen issues. Weaker responses often attempted many evaluation issues without discussing any of them in depth or giving any examples from the **part (a)** of the response to back up their points. These types of responses often achieved Level 1.

Psychology and Health

Question 5

(a) This was often well answered with many responses achieving full marks with good outline of one symptom of Munchausen syndrome. The most common response identified the disorder as faking an illness in order to gain attention. Many responses gave examples of how the patient would fake symptoms such as tampering with test results or self-harm. Weaker responses tended to be very brief and just identified that the syndrome involved faking or lying about an illness. Some responses outlined hypochondriasis or gave symptoms such as pain and fatigue. These types of responses were not creditworthy.

(b) There were many strong responses to this question with many giving a detailed outline of findings from the study by McKinstry and Wang on non-verbal communications in the patient–practitioner relationship. Common findings included preference for a formally dressed male doctor in a suit, a female doctor in a lab coat and older and higher social class participants showing a strong preference for more formal attire. Weaker responses gave fewer findings from the study. In addition, many responses outlined the procedure of the study which was not creditworthy.

(c) Most responses identified at least one weaknesses, if not two, of the McKinstry and Wang study. A common weakness given was the lack of generalisability as participants were from Scotland, and that these findings may not refer to other countries. Stronger responses were able to clearly explain why participants in another country might have a different response to the participants in the McKinstry and Wang study, and this was sometimes backed up with a clear example. Another well-described weakness was the failure to include a picture of a woman in a suit. Some responses outlined a weakness of the self-report used such as collecting quantitative data or social desirability bias. These types of responses were often weaker as they were not well explained with clear examples from the study.

Question 6

(a) Responses to this question covered the full range of the mark scheme. Stronger responses had a good understanding of the research into improving adherence to medical advice. Some responses gave clear and accurate details of how to improve practitioner style as outlined by Ley and the research by Yokley and Glenwick, and Watt et al. Weaker responses often outlined anecdotal ideas about how to improve adherence and gave fewer details of the research, or some of the details were incorrect. These types of responses often achieved in Level 1 or Level 2. Some responses outlined measurements of adherence to medical advice, such as pill counting or repeat prescriptions, but did not make any reference to improving adherence. These types of responses were not creditworthy.

(b) There were some strong responses to this question. These were often able to evaluate the named issue of experiments and discussed issues such as the experimental design and ecological validity,

and could give some examples from the Yokley and Glenwick and Watt et al. studies described in **part (a)**. A few attempted some analysis and were able to discuss both the strengths and weaknesses of the experimental method and apply these issues effectively to the research. Other common evaluation issues included generalisability, practical applications, reliability and ethics. Weaker responses often did not cover the named issue or outlined strengths and/or weaknesses of experiments without any reference to the research from **part (a)**. Some responses outlined a number of issues but in a superficial manner, simply stating that issues did or did not apply to studies but not why. These types of responses often achieved in the Level 1 mark band.

Psychology and Organisations

Question 7

(a) There were some good responses to this question with some achieving full marks by giving a clear outline of the two-factor theory of job satisfaction. Strong responses outlined how satisfaction and dissatisfaction at work are independent of each other and mentioned motivators and hygiene factors with an example of each.

Weaker responses often outlined motivators and hygiene factors but were not able to outline how these link to satisfaction/dissatisfaction or that they are independent of each other. These types of responses often achieved 1 mark.

(b) Most responses were able to achieve 2–3 marks by describing the five dimensions measured in the job descriptive index (JDI) and how it is scored. A few responses were able to give an example of one of the statements used in the JDI. Weaker responses gave fewer details of the JDI and often listed the five dimensions and no other details. A number of responses gave incorrect details of the measure and often stated that this is used to determine whether a candidate should be given a job, rather than being a measure of satisfaction at work. These types of responses were not creditworthy.

(c) Responses to this question covered the full range of the mark scheme. Stronger responses outlined both a similarity and a difference between the JDI and the quality of working life (QWL) questionnaire. Most common good similarities explained how both measures can be useful to organisations and how they could be used, or that both collected quantitative data with examples given from both measures. Well explained differences included an outline of the number of dimensions measured in each, with examples given or the rating scales used in the measures. Weaker responses often identified the similarity or difference without providing any examples. Some responses suggested that one of the measures used qualitative data which is incorrect and therefore not creditworthy.

Question 8

(a) There were some good, detailed responses to this question with clear details given of group conflict in organisations, including details of the levels and causes of group conflict, positive and negative effects of conflict and how to manage group conflict from Thomas. Weaker responses frequently outlined one of the bullet points from the section and therefore achieved in the lower mark bands. There were also some responses that gave anecdotal responses outlining causes of group conflict in an organisation. In addition, some responses outlined theories from other parts of the syllabus such as group development and roles without any reference to group conflict. These types of responses were not creditworthy.

(b) There were some strong responses to this question where it was structured by evaluation issue and began with the named issue of individual and situational explanations. Some of these responses gave clear examples from **part (a)** of their response to back up their evaluative points and attempted some analysis. In addition to individual and situational explanations, popular evaluation issues were generalisability and practical applications. A number of responses often continued the description from **part (a)**, and did not provide any evaluative points which was not creditworthy.

PSYCHOLOGY

Paper 9990/33
Specialist Options: Theory

Key messages

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

It is important that candidates are made aware of the terminology, concepts and theories identified in the syllabus, as well as key terms used in named theories and studies, as some were unable to identify and/or define the terms given in these types of questions. Creating a glossary of key terms, revision of terminology using flash cards and class quizzes on terminology could prove useful. Where the response gave an example to help define the term or theory, it often achieved full marks. These questions are worth 2 marks and a brief response is appropriate.

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

These questions could ask the candidate to describe a theory, study, treatment or technique 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 the findings, or a summary of the key features of the study. This question is worth 4 marks and the candidates should write a more extended answer. It would be helpful for candidates to create a revision flashcard or mind map of each bullet point in the syllabus. The flashcard should be given the title used in the syllabus, for example, 'treatment and management of schizophrenia and delusion disorder: cognitive-behavioural therapy: Sensky', to help the candidate to identify which part of the syllabus the question is referring to. If the question asks for a part of the study, the response should only describe this part. For example, if the question asks for the procedure, then the response should not describe the findings of the study.

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

These questions could require the candidate to explain up to two strengths/advantages or weaknesses/ disadvantages 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 such as the effectiveness of controls used in a study. 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 named in the question. To improve, responses should give specific examples to support their point. As mentioned for the odd question **part (b)**, the candidate could make a flashcard/revision notes and include strengths and weaknesses of the theory, study or technique to help candidates prepare for these questions.

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

This question comes from one of the bullet points in the syllabus. Candidates could describe the three or four studies, theories or techniques identified in the syllabus under the appropriate bullet point. It is possible for the responses to achieve full marks by describing at least two of the studies, theories or techniques and this would need to be a very detailed description. Full marks can also be achieved by a response that describes three of the bullet points in detail (in less depth than if the response described two of the studies, theories or techniques) with excellent understanding and good use of terminology throughout. It is also important that the descriptions are linked to the topic area named in the syllabus. It could be useful for candidates to create revision notes with the title of each bullet point as the header in their notes.

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

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

Some of the responses covered other issues rather than the one named in the question. Quite a few of the answers were structured by study/theory/technique rather than by the issue, which often led the response to be quite superficial and repetitive. A few of the responses did analysis. Candidates should be aware this question is worth 10 marks and attempt to include an appropriate amount of information.

General comments

The marks achieved by candidates for this session of the 9990 syllabus achieved across the full range of the mark bands.. Some candidates were well prepared for the exam and showed good knowledge, understanding and evaluation throughout their responses. A number of candidates were not as well prepared and showed limited knowledge and understanding with brief and/or superficial responses. These candidates often showed limited evaluation skills.

Time management for this paper was good for the majority of candidates and most attempted all questions that were required. Some candidates did not respond to one or more of the questions asked in the option area. A very small number of the candidates attempted to respond to more than two topic areas but often did not attempt all 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) Many of the responses addressed the question and achieved one or two marks by outlining a cognitive explanation of schizophrenia and delusional disorder. Common responses included outlining faulty reasoning, difficulties with self-monitoring, and a failure to recognise hallucinations as that of the sufferer's own voice. Some responses outlined a behavioural or biochemical explanation of schizophrenia which was not creditworthy.

(b) A significant number of responses were able to describe the procedure in the study by Sensky et al. of cognitive-behavioural therapy (CBT) for schizophrenia. Strong responses referred to the participants, the random allocation to treatment, details of CBT and befriending treatments, as well as some details of how the patients were assessed. Weaker responses gave an outline of CBT with no other details of the study provided. Some responses described the incorrect study such as the study on token economies by Paul and Lentz. These types of responses were not creditworthy.

(c) The responses to this question covered the full range of the mark scheme. Stronger responses outlined both a similarity and a difference between CBT and a biochemical treatment for schizophrenia/delusional disorder. Most common good similarities looked at intervention from therapist or evidence for success, good differences included the fact that the treatments were

based upon different approaches and prevalence of side effects in drug treatments but not CBT. Weaker responses often identified the similarity or difference without providing any examples. Some responses compared CBT to ECT which is not a biochemical treatment. These types of responses were not creditworthy.

Question 2

(a) Responses varied considerably for this question and covered the full range of the marks available, although most achieved in the lower mark bands due to lack of detail in the response and/or errors in the description.

Stronger responses gave good definitions of bipolar and unipolar disorder as well as detailed outline of the BDI. Weaker responses described bipolar disorder but stated that the change from depression to mania could happen very quickly and last a very brief period of time which is incorrect. There were a number of descriptions of explanations and/or treatments and also a significant proportion of candidates who described a number of disorders from other parts of the syllabus, notably schizophrenia, phobias and OCD. These types of responses were not creditworthy.

(b) Many of the responses achieved in the Level 1 or Level 2 mark band. A few candidates provided clear analysis and details of the characteristics and measures to back up their evaluative points that enabled their responses to achieve Level 3 and above. There was a tendency for responses to focus on many issues per characteristic/measure rather than applying the issue to the different characteristic/measure.

A few responses did effectively discuss the named issue of validity with some clear analysis. Many of the responses did not address the named issue of validity. Common evaluation issues included self-reports, reliability and application to everyday life. Some responses just evaluated the BDI with no attempt made to evaluate the characteristics and the issues surrounding diagnosis. Some responses evaluated explanations and/or treatments in **part (b)** which was not creditworthy.

Psychology and Consumer Behaviour

Question 3

(a) There were a few strong responses to this question with an identification of two stages of the consumer decision model other than 'recognition of need'. Popular stages included identifying 'informational search' and 'evaluation of alternatives'. Weaker responses identified one stage correctly. Many candidates did not attempt this question or gave incorrect stages.

(b) There were some very clear and detailed responses describing two different ways to close a sale. Common responses included reciprocity, liking, scarcity and body language. Responses that gave examples to explain how the way to close a sale would lead to a purchase often achieved full marks. There were some candidates that did not attempt this question.

(c) There were some strong responses to this question with the vast majority being able to give both an advantage and a disadvantage of one of the ways to close a sale outlined in **part (b)**. Common advantages included explaining why this way of closing the sale would lead to a purchase and increased income for the company, repeated sales in the future through word of mouth and increased reputation of the company/product. Common disadvantages included explaining why the consumer might not purchase the product due to feeling pressured and lowered reputation of the company. Strong responses identified the advantage/disadvantage and gave clear examples to explain their point. Weaker responses did not explain with any clarity about why it was a strength or weakness. For example, many simply outlined that the company would have increased sales rather than looking at the point of view of the consumer and their satisfaction. These types of responses achieved in the Level 1 mark band.

Question 4

(a) Many responses achieved Level 1 or Level 2 for this question. A few were able to achieve in Level 3 or Level 4 by giving fairly detailed descriptions of the models of the effects of ambience including pleasure-arousal and cognition-emotion, the Kutlu et al. study on lighting and colour in retail stores and the Chebat and Michon study on the effects of odour on shopper arousal and emotion. Weaker

responses often briefly described the studies by Kutlu et al. and the study by Chebat and Michon but often with few details. Some responses just outlined the findings/conclusion of these studies with no other details given. These types of responses often achieved in the Level 1 mark band.

(b) Those responses that achieved in the higher mark bands for **part (a)**, tended to produce good answers to this question, with some understanding of how control of variables applied to the studies with some examples given. Some attempted analysis, for example pointing out that samples in both the studies lacked generalisability with some examples given to back up the point made. Weaker responses did not attempt the named issue of control of variables. This appeared to be because as the study was not described in **part (a)**, the candidate was not aware of the controls used in the studies. Ecological validity, ethics and self-reports were other commonly chosen issues. Weak responses often named an issue and then stated that a study either supported the issue (had good ethics) or did not support the issue (had poor ethics) without any explanation given for the point made. These types of responses achieved in the Level 1 mark band.

Psychology and Health

Question 5

(a) Many responses achieved 1 mark by giving examples of the features/components of the Yale model of communication such as the communicator, target, attention, etc. However, most responses did not outline the model in relation to health promotion to achieve full marks. A few responses did this by giving an example related to health promotion for one of the features, such as the communicator being a well-respected doctor/expert in the field of medicine, and achieved full marks. A number of responses did not know the Yale model of communication and either did not answer the question or wrote a general response about health. These types of responses were not creditworthy.

(b) There were a few strong responses to this question with some giving a detailed description of the Lau et al. study on health change in adolescents. Common responses included the aim, that the study was longitudinal, that adolescents and their parents answered questionnaires and some of the results of the study. Weaker responses often outlined that the study used a questionnaire and gave a result. There were many incorrect responses with candidates outlining a study done on children rather than adolescents. These types of responses were not creditworthy.

(c) Where there was a creditworthy response in **part (b)**, the candidates were often able to outline both a strength and a weakness of the Lau et al. study. Common strengths included generalisability of the sample size, practical application and being able to monitor change in health belief/change over time. The most common weakness was generalisability due to the participants being from one university. Many responses achieved Level 1 or Level 2 due to not providing any clear example from the Lau et al. study to back up their strength and/or weakness. Responses that were not creditworthy typically evaluated the incorrect study.

Question 6

(a) The responses to this question covered the full range of the mark scheme. Stronger responses had a good understanding of the measures of pain and focused on 3 or 4 measures rather than attempting to describe all of the measures detailed on the syllabus. There were some very detailed descriptions of the McGill Pain questionnaire, UAB pain behavioural scale and the Wong-Baker scale. Weaker responses attempted to describe a large number of measures of pain but very superficially (and/or incorrectly). These types of responses typically achieved in the Level 1 mark band. A number of responses described theories of pain rather than measures of pain which was not creditworthy.

(b) There were some strong responses to this question. These were often able to evaluate the named issue of psychometrics and give the strengths and weaknesses of this type of measure, as well as examples from the measures of pain described in **part (a)**. Other common evaluation issues for this question included the reliability of the measurements, self-reports (although this sometimes just repeated the issues of reliability and/or validity), quantitative data and objectivity versus subjectivity. Weaker responses seemed to be when a large number of issues were covered but all were done in a very superficial manner, simply stating that issues did or did not apply to the measurements but not why. These types of responses often achieved Level 1.

Psychology and Organisations

Question 7

(a) There were some strong responses to this question with some achieving full marks. The most common practices identified from Kouzes and Posner's Leadership Practices Inventory (LPI) were 'encourage the heart' and 'model the way' which was either identified or described. Weaker responses identified one of the practices (typically 'model the way'). Some responses gave an outline of a theory of leadership which was not creditworthy.

(b) There were several good, detailed responses to this question. There were some strong answers describing the styles of leader behaviour proposed by Muczyk and Reimann including the dimensions of directive/permissive and autocrat/democrat. Other good responses described specific types e.g. directive autocrat. Some responses confused the styles of leader behaviour with situational leadership. These types of responses were not creditworthy.

(c) There were a few good responses to this question with both a similarity and a difference given between Muczyk and Reimann's styles of leader behaviour and Hersey and Blanchard's situational leadership. Stronger responses outlined the similarity that both depend on the situation or require the leader to be flexible/adaptable, and some did give examples from both theories to back up their similarity. A common difference was that the styles of leader behaviour sees a distinction between making decisions and carrying them through whereas situational leadership does not. If the candidate outlined situational leadership in **part (b)** of their response, they were often unable to outline either a similarity or a difference.

Question 8

(a) There were some good, detailed responses to this question with clear details given of group development and roles in organisations, including Tuckman's theory of group development, Belbin's theory of team roles and Belbin's team inventory. Tuckman and Belbin's theories tended to be the most detailed of these descriptions. Weaker responses tended to be brief or gave anecdotal descriptions of group development and roles. A number of responses referred to other sections of the syllabus, such as attitudes to work and theories of motivation. These types of responses were not creditworthy.

(b) There were a few good responses to this question where it was structured by evaluation issue and began with the named issue of practical applications. Some of these responses gave clear examples from **part (a)** of their response to back up their evaluative points and attempted some analysis. In addition to practical applications, popular evaluation issues were individual and situational, and reductionism versus holism. Weaker responses did attempt to evaluate the named issue but did not explain why the research had a practical application and did no analysis. These types of responses tended to state that a theory or study either was or was not applicable to companies with no explanation of this given. Those responses that attempted individual and situational explanations were often weak with confusion over the definition of this issue.

PSYCHOLOGY

Paper 9990/41
Specialist Options: Application

Key messages

- What has been learned from the AS component of the syllabus should be transferred to the A2 component. For example, at AS candidates learn about methodology, such as experiments, which also applies to A2.
- Questions should be read carefully, ensuring that the focus is on what the question asks.
- 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) and a conclusion.
- 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.
- 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.
- Psychological knowledge should be applied wherever possible. Anecdotal and common-sense answers will not achieve top marks.

General comments

There was evidence to suggest that many candidates had not studied two options to the same level. Whilst answers to one option were often very strong, some answers to the second option were notably weaker, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

- 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.
- Answers must refer to the study the question is about. Many answers provided general comments which were unrelated to the study itself.
- For **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so restricted marks.
- Many conclusions merely repeated what had already been written, and such summaries scored no marks. A conclusion is a 'decision reached by reasoning' and so as the reasoning has been done through the advantages and disadvantages, a final decision/conclusion needs to be drawn.
- Candidates should ensure they are focusing on 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.
- Candidates should provide sufficient detail to score all the available marks. A single sentence is more likely to score one mark rather than two marks, so a little elaboration, explanation or an example that goes beyond the basic sentence is always recommended. Candidates should always try to include psychological knowledge in their responses.

Section B

Answers to **part (a)** questions in this section should include an appropriate design and 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. Describing a relevant piece of research from a topic area alone is not enough to score full marks. The links between the research and how it informed the design must be shown. There is no need to quote a name (date) for each sentence, some responses included 'I chose a self-selecting sample because Milgram (1963) did' for example. This just identifies a study using that technique, rather than explaining 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. Therefore, candidates are not required to describe everything they know about that topic area, and answers that do not address the question will only achieve minimal marks. To score marks at the top end of the mark range, answers must focus on arguments both for and against the statement, and then 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) Many candidates were able to clearly explain the term 'covert sensitisation' and were awarded full marks. A few candidates confused it with imaginal desensitisation, systematic desensitisation and even EMDR. No marks were awarded for incorrect explanation of terms.

(b) (i) Some very strong answers showed good understanding of relevant terms and explained how they were applied in the study by Glover. However, there were also some weak answers which with little knowledge demonstrated about the Glover study. An example of an incorrect response was suggesting that covert means the participant does not know they are being treated.

(c) Many candidates appeared to confuse the terms or did not show understanding of the difference between the terms sensitisation and desensitisation or between imaginal and covert. A few candidates described pictures being shown to a participant, which is incorrect, or defining covert as the participant not knowing that they are involved in treatment, also incorrect. Answers scoring very few marks also described one term followed by the other rather than giving differences as the question required. One difference is that imaginal desensitisation uses progressive muscle relaxation whereas covert sensitisation does not. Another difference is that covert sensitisation creates an unpleasant association (vomiting in the Glover study) whereas imaginal desensitisation reduces the emotions associated with an event.

(d) All **part (d)** questions in ***Section A*** require a discussion of advantages and disadvantages and a conclusion, and this question part was no exception. Further, the advantages and disadvantages must be related to the question, in this instance covert sensitisation. There were a small number of answers achieving full marks, with many candidates not relating their answer to the topic area. Summaries were often included rather than conclusions.

Question 2

(a) The important term that was essential to be addressed in this question was the term 'perceived' because this means that the experience of crowding can vary from one individual to another. Many candidates included this and scored full marks. Others did not show sufficient understanding of the term.

(b) (i) Answers to this question were frequently awarded full marks for stating the difference between the two terms. Candidates understood that human density involves people whereas spatial density is about the physical location, and with regard to consumer psychology, the number of fixtures and fittings of a store, or the amount of merchandise in the store, for example.

(ii) This question asked about the questionnaire used by Machleit et al. to gather data about perceived crowding. Many candidates gave non-specific and general answers such as 'it included a number of questions'. The questionnaire used a Likert-type scale, scale range 1–7, asking questions like 'the store seemed very crowded to me' each of which would score 1 mark.

(c) Many candidates were unable to suggest a model of the effects of ambience and this restricted their marks. Some candidates suggested models of personal space, which also could not be awarded marks. Some candidates suggested the cognition-emotion model, which was applicable, but the strongest answers suggested the Mehrabian and Russell PAD model which is directly applicable. This model considers the levels of Arousal, Pleasure and Dominance (the amount of control a person feels over a situation). This model, applied to the effects of crowding, earned some candidates maximum marks.

(d) Many answers included two strengths and two weaknesses of questionnaires, but only scored partial marks because, frequently, answers were not related to the effects of crowding (or even consumer behaviour) as the question required. An answer might be 'one strength (of using a questionnaire) is that it produces quantitative data which can be statistically analysed'. This is correct but it would not receive full credit without being related to the question, because it could apply to any question from any topic area rather than this specific question.

Question 3

(a) All candidates had some knowledge about the Funhaler device and all scored at least one mark. Answers whose description was more detailed, such as adding more than one point, or that it included a whistle and spinner not included in other devices, were awarded full marks.

(b) At the top end of the mark range, answers were referring to positive reinforcement, Skinner and operant conditioning, and often explained fully how the spinner and whistle were reinforcers. There were some weak answers at the bottom end of the mark range suggesting nothing more than it is a behavioural device because it involves behaviour.

(c) Answers which were not outlined by Watt et al. were acceptable, but candidates suggesting these often failed to provide any more detail than a single sentence. Although this earned one mark (for each suggestion), answers needed detail to elaborate to gain two marks. For example, a one mark answer might be 'children don't take their medicine because they do not like the taste' whereas a two mark answer might be 'children apply rational non-adherence: they do not like the taste of the medicine so make the logical decision that they would rather not take the medicine than suffer the bad taste.'

(d) Many candidates scored no marks or very few marks because they did not answer the question set. The question was specifically 'studies on non-adherence in children' but responses were often a general discussion on using children in psychological studies, using AS level studies rather than those specifically relating to adherence. Any advantage or disadvantage must be focussed on the question in order to achieve marks.

Question 4

(a) There were some excellent explanations of the quality of working life (QWL), most referring to QWL as 'satisfaction with personal and working needs'. Alternatively, there were answers which merely stated that quality of working life is about the quality a person has in their life at work. Repeating or rewording the question cannot be credited.

(b) An outline of two features used to assess QWL were required. Some candidates were not able to demonstrate any knowledge, but others identified a feature (one mark) and then provided an outline to explain the feature (two marks). A maximum mark answer was typically: 'One feature is social integration (one mark) which refers to the interpersonal relationships a worker has with other workers at work' (two marks).

(c) (i) A small number of candidates confused reliability with validity, which couldn't be credited. Some answers were brief, with nothing more written than 'use test-retest' for limited credit. Others wrote 'give the questionnaire to other workers' which is incorrect because to test reliability, the *same* questionnaire should be given to the *same* workers at a later date (i.e. test re-test).

(ii) Many candidates did not make a suggestion and some suggestions were vague, but top marks went to those who suggested using concurrent validity, comparing the results of the QWL test with some other measure of satisfaction. Other measures such as the job descriptive index (JDI) or the Minnesota Satisfaction Questionnaire were appropriate.

(d) This question part followed on from the theme introduced in the stem and asked for advantages and disadvantages of using a five-point scale to measure job satisfaction. Relevant advantages and disadvantages were given, but often these were not always related to the QWL questionnaire (or any other measure). Any advantage or disadvantage must be focussed on the question in order to achieve marks.

Section B

Question 5

(a) This question did not receive many strong responses . The question asked for an investigation into the long-term effectiveness of applied tension. A few candidates knew that applied tension is used to treat blood/injection phobia and planned very good studies. The main weakness was that many candidates did not know what applied tension is used for. It was commonly assumed that it is a treatment for phobias and so a study was planned comparing applied tension with systematic desensitisation (or some other way to reduce anxiety). The problem is that applied tension raises blood pressure whereas other techniques lower blood pressure through relaxation.

(b) The psychological evidence included in this question part needed to focus specifically on applied tension and blood/injection phobia. Whilst a few candidates did this, many candidates wrote about many different kinds of phobias, which was not relevant to the question.

Question 6

(a) Most candidates opted to plan a field experiment, which was a logical choice given the question. Designs were often impressive because the complex IV had to include three conditions of chocolate products placed in three different ways in a film clip. Some opted to place a product at the beginning, in the middle or at the end of a film clip. Others chose to have the product held in the hand, on a table, or an alternative location. The DV was often the number of times the product was noticed when participants were asked in a later questionnaire. Overall, there were some very strong answers.

(b) Methodologically some answers were very strong with detailed explanations for the choice of placement, and many answers also focused on the controls that were applied. Some answers needed to explain the plan in more detail, rather than a single sentence statement. Such answers listed things like 'I used an opportunity sample' without any further comment, but an explanation is required. Psychological knowledge was very strong in many answers, with candidates often explaining how the study on product placement by Auty and Lewis related to their plan.

Question 7

(a) This question on age differences in non-adherence allowed candidates a free choice of method with many choosing to make the study an experiment. However, the consequence of this was that the IV was often muddled between different ages and different reasons. The easier choice, as done by a small number of candidates, was the use of a questionnaire (or an interview), which asked participants of different ages to choose one reason for their non-adherence from a range of different possibilities. Taking time to think about the most appropriate method to use and thinking the plan through before starting to write is time well spent. It is also worth noting that an experiment cannot automatically be applied easily to every question that is set.

(b) Direction to what psychological evidence to use to investigate reasons for non-adherence was provided in the stem of the question, which stated: 'the health belief model provides many reasons

to explain non-adherence'. Yet most candidates overlooked this and invented reasons of their own. These reasons were usually not based on psychological knowledge. A few candidates opted to write about rational non-adherence and this was creditworthy. Methodologically most experiments were not well planned with confused IVs. The methodology for questionnaires was often good, although the reasons for choosing to use open-ended questions for example was often insufficiently detailed.

Question 8

(a) This question also gave a free choice of method, and again an experiment was chosen to be applied by most candidates. Most of these candidates chose to have an IV of 'with targets' and 'without targets' and a DV of 'level of motivation' which was good. What was lacking was knowledge of target setting, such as the five features of SMART targets, with many candidates referring to no more than 'target setting' throughout. The plan should be based on the psychological knowledge explained in **part (b)**. If an experiment is chosen, it is important to include all the specific features (type, IV, DV, controls, design) and also to include a number of general features, such as the sample and sampling technique, the type of data to be gathered and appropriate ethical guidelines.

(b) Methodologically most answers were good but could have been made better with more explanation of the reasons for design decisions. The psychological evidence should have focused on Latham and Locke's goal setting theory and their SMART targets (where targets should be specific, measurable, attainable/agreed, realistic/relevant and time-based). A few answers focused on this work but many did not, instead focusing on the word 'motivation' and writing about Maslow and his hierarchy which was not relevant.

Section C

Question 9

The main debate here was whether the psychoanalytic explanation of phobias could be generalised to everyone, with an additional debate about case studies. There were three types of answers. Firstly, some candidates focused exclusively on the study of Charles, and although advantages and disadvantages of case studies were provided, the psychoanalytic component was absent. Secondly, there was a debate about phobias in general, with reference to little Albert and classical conditioning. Again, there was an absence of the psychoanalytic approach. The third type did include the work of Freud, but often did not go beyond the work of little Hans. Only a few candidates realised that the Psychoanalytic/Freudian explanation of phobias applies to all, suggesting that phobias are due to the id, ego and superego.

Question 10

There were some strong answers in response to this question, showing that candidates knew about eye-tracking, a quantitative, 'scientific' measure that can be used when studying menu item position. The debate was whether this alone was sufficient without the need for qualitative data, or whether people should be asked about what they look at and why when studying a menu. Some detailed answers addressed this debate, but others were superficial, often making vague points that were unrelated to menu design.

Question 11

This question linked unrealistic optimism, outlined by Weinstein with preventing stress. The debate would logically be that those who are unrealistically optimistic would not need to prevent stress, and those that are not unrealistically optimistic would consider doing something to prevent stress. Whilst a few candidates were able to debate this, others struggled due to lack of knowledge about unrealistic optimism, preventing stress or both.

Question 12

The words 'need to achieve' in the question should have focused minds on the syllabus section of achievement motivation and the work of McClelland (1965). McClelland argues that a need for achievement is very important in working life, as is the need for affiliation (to work with and enjoy friendships with work colleagues). Consequently, earning money is less relevant. However, many candidates did not address this and focused on Maslow and his hierarchy of needs.

PSYCHOLOGY

Paper 9990/42
Specialist Options: Application

Key messages

- What has been learned from the AS component of the syllabus should be transferred to the A2 component. For example, at AS candidates learn about methodology, such as experiments, which also applies to A2.
- Questions should be read carefully ensuring that the focus is on what the question asks.
- 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) and a conclusion.
- 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.
- 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.
- Psychological knowledge should be applied wherever possible. Anecdotal and common-sense answers will not achieve top marks.

General comments

There was evidence to suggest that many candidates had not studied two options to the same level. Whilst answers to one option were often very strong, some answers to the second option were notably weaker, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

- 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.
- Answers must refer to the study the question is about. Many answers provided general comments which were unrelated to the study itself.
- For question **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so restricted marks. For example, to score one mark, answers must include an advantage and an example.
- Many conclusions merely repeated what had already been written, and such summaries scored no marks. A conclusion is a 'decision reached by reasoning' and so as the reasoning has been done through the advantages and disadvantages, a final decision/conclusion needs to be drawn.
- Candidates should ensure they are focusing on 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.
- Candidates should provide sufficient detail to score all the available marks. A single sentence is more likely to score one mark rather than two marks, so a little elaboration, explanation or an example that goes beyond the basic sentence is always recommended. Candidates should always try to include psychological knowledge in their responses.

Section B

Answers to **part (a)** questions in this section should include an appropriate design, and 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. Describing a relevant piece of research from a topic area alone is not enough to score full marks. The links between the research and how it informed the design must be shown. There is no need to quote a name (date) for each sentence, some responses included 'I chose a self-selecting sample because Milgram (1963) did' for example. This just identifies a study using that technique, rather than explaining 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. Therefore, candidates are not required to describe everything they know about that topic area, and answers that do not address the question will only achieve minimal marks. To score marks at the top end of the mark range, answers must focus on arguments both for and against the statement, and then 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) In response to this question part, answers were very general rather than based on the work of Ellis. Ellis believed that, as one candidate wrote 'a person is not directly affected by outside things but by their own perception of external things' and so for Ellis 'B: Beliefs about the event' are crucial. It is the beliefs about an event that therefore need to be changed.

(b) (i) Only a very small number of answers included an irrational belief outlined by EllisExamples of irrational beliefs outlined by Ellis include 'awfulising', 'self-downing', musterbating' and 'I can't stand it itis'. Awfulising, for example, is when a person thinks 'if I don't succeed it will be awful'.

(ii) Many answers outlined how irrational beliefs could be changed, but often answers were not based on REBT, instead a general comment about changing thinking was provided. Ellis outlines D: the Dispute phase, E: the new more effective Emotions, and Acceptance where the person accepts reality, whether it is pleasant or unpleasant. Answers referring to some parts of this process were awarded marks.

(c) This question asked for an outline of a *psychological* treatment of depression. Both Ellis and Beck provide psychological treatments and therefore candidates outlining the treatment for depression outlined by Beck were awarded marks. A few candidates described Beck's *explanation* of depression (e.g. the cognitive triad), rather than *treatment,* which could not be credited. Some candidates wrote about electroconvulsive therapy (ECT) and drug treatments, but as these are *biological* rather than *psychological*, and could not be credited.

(d) This question part continued with the *psychological* treatment theme and so answers based on biological (or medical) treatments only received marks if they were used in discussion about psychological treatments. Some candidates wrote excellent answers and achieved maximum marks, but many answers only scored partial marks because the strengths and weaknesses were not supported with appropriate examples from either Beck or Ellis.

Question 2

(a) Most candidates were able to score full marks by stating that the Kranes design was preferred (one mark) because it yielded significantly higher ratings than the Friedman design on pleasure and restoration (2 marks). A few candidates confused the two designs and scored no marks.

(b) The question required two differences; it did not require a description of one design followed by the other. Candidates scoring full marks stated clearly that, for example, 'the Kranes design has high ceilings whereas the Friedman design has low ceilings' (2 marks). Another correct answer would be that the Kranes design aims to maximise pleasure and restoration, whereas the Friedman design decreases restoration.

(c) (i) Any two features of the sample were creditworthy and many candidates scored full marks. Relevant features included: 48 participants, 26 males and 22 females, or 22 participants who provided ratings. Additionally: participants had gambled in all six casinos; the casinos were in Las Vegas, United States; the mean age was 28 years; all had at least a high school degree.

(ii) Most candidates were able to give a disadvantage of using opportunity sampling and scored one mark. Some answers were no more than one sentence, but many others gave more detail or related it to the sample in the Finlay et al. study and gained an additional mark.

(d) Many answers included two strengths and two weaknesses and a conclusion, but only scored partial marks because often answers were not related to the findings of the Finlay et al. study as the question required.

Question 3

(a) This question required two findings from the graph (shown in fig 3.1). Most candidates answered the question clearly, often quoting numbers to support their comment. For example: 'one finding is that heart rate is the highest in the high stress task reaching approx. 85 beats per minute'. Answers like this (with a second finding written in this way) scored maximum marks. Other candidates gave vague answers such as 'it is highest in the high stress task and low elsewhere' for limited credit.

(b) (i) Although some candidates outlined the two stress tasks used in the Wang et al. study correctly, many others were not able to. In the Wang et al. study there was a low stress task, counting aloud backwards from 1000 and a high stress task, performing serial subtraction of 13 from a four-digit number out loud.

(ii) This question required an explanation of what causes an increase in heart rate when stressed. Answers ranged from those which referred to no physiology at all, to those where knowledge and understanding were well applied. Most answers referred to the pituitary gland secreting hormones, ACTH stimulating the adrenal medulla to produce adrenaline which causes an increase in heart rate. A few answers referred to the similar process involving cortisol production.

(c) There were some very strong answers which scored full marks and many incorrect answers. In the Wang et al. study there were four measures used to test validity: two physiological measures, those of heart rate and salivary cortisol, and two psychological measures (i.e. the question) which were a self-report of stress and anxiety (scale 1–9) before/after each task and also a self-report of effort, frustration and task difficulty (scale 1–9) only after stress tasks.

(d) This part followed on from the theme introduced in the stem regarding physiological measures. Many answers were restricted because of a lack of knowledge about the physiological measures used in the Wang et al. study, even though the stem included details about the use of heart rate. As previously mentioned, the stem of the question provides essential information on which all question parts are based. Marks were also limited because many candidates provided a summary of what already been written rather than a conclusion.

Question 4

(a) Correct answers, frequently achieving full marks, stated that temporal is time-related and often gave a supporting example by referring to shiftwork. Consequently, many answers scored no marks as they confused temporal with temperature, working conditions or temporary employment. The stem of the question is crucial to the questions that follow, and reading this would have directed candidates to the link between temporal and time.

(b) (i) Nearly all candidates were able to give a correct difference between the metropolitan and in most cases, the continental rota. The difference is that the metropolitan rota is based on a 2-day cycle whereas the continental rota includes both 2-day and 3-day work periods.

(ii) This question part also resulted in most candidates scoring full marks for suggesting two correct effects of shiftwork on health. Most answers mentioned peptic ulcers (or gastrointestinal problems), cardiovascular disease, and pregnancy difficulties. Some answers correctly referred to sleep problems. A few answers claimed that there is increased risk of cancer. However, Knuttson (2003) states 'there is no evidence for increased risk of cancer.'

(c) Many candidates could not suggest two ways in which working hours can be organised, other than by shiftwork. There are many possibilities. Working a 9–5 'standard' day is very common across the world. Variations of this are also possible. Some workers have 'flexi-time' where they choose their own working hours provided 8 hours per day are worked, or variations on this, a 'compressed' work week, allowing working more hours per day giving more days off work. There is an 'on-call' system which is commonly used by medical staff in hospitals.

(d) Some good answers were written for this question part on shiftwork, but many suffered from the same problems outlined above for **part (d)** and candidates could not be awarded marks. In addition, for this question, even though the question stated 'other than in relation to health' many answers included health in their answers and so marks could not be awarded. Candidates must answer the question that is set to allow them to be awarded all the available marks.

Section B

Question 5

(a) Overall, this question did not receive many strong answers. The question asked for an investigation into the most common type of compulsive behaviour. The most logical plan would be to design a questionnaire listing possible compulsive behaviours, asking participants to tick all the behaviours they did and then analysing to see which was the most common. Instead of this, many experiments were conducted comparing males and females, different types of OCD, and many studies were designed to test different types of treatment.

(b) The psychological evidence included in this question part should have focused specifically on *compulsive* behaviors, such as checking, cleaning/washing, slowness, doubting, ordering, repeating. These were frequently absent from many answers. Instead, the study of 'Charles' was described in full, with little focus on his compulsive behaviour. Methodological knowledge was sometimes good, but if the plan of the study was flawed, such as those answers focusing on treatments, then the methodology used to test this was also flawed (an independent design comparing two different treatments, for example).

Question 6

(a) All candidates answered the question and planned a field experiment. Most IVs consisted of comparing one or more different odours or comparing an odour with a 'no-odour' control. Often there was good detail in the procedure and DVs were often logical, such as recording how long people stayed in the clothes shop, or the number of purchases made in the different IV conditions. Some suggested using a structured interview to ask shoppers whether they had noticed the odour or not. Overall, there were some very good plans suggested.

(b) Psychological knowledge was described competently in many answers, with candidates focusing on the study by Chebat and Michon (2003). However, some answers did not go on to explain how this study informed their plan and so could not be awarded full marks. Methodology was often very good, with a range of specific and general methodological features being included.

Question 7

(a) This question allowed candidates a choice of experiment and most opted for a field experiment. Plans were often very good, with variables clearly designed. Most candidates opted to give a drug to one group and an alternative technique to a different group. However, some candidates were confused about the purpose of alternative techniques, and some did not know the difference between acupuncture and TENS. This resulted in some ambiguous procedures being suggested.

(b) Psychological knowledge was sometimes vague because some candidates knew very little about acupuncture and stimulation therapy (TENS). Other answers showed a detailed knowledge and good understanding and often answers linked this knowledge with the plan of the study. Methodological knowledge was sometimes very good, but again those with limited knowledge had confused IVs and procedures.

Question 8

(a) Answers to this question resulted in limited marks through not addressing the question set. The question required a focus on the 'styles of leader behaviour shown by a person in charge of a school'. This meant that any plan involving a range of leaders, in a number of different schools, or a plan involving different organisations could not be awarded full marks. A second problem was that Muczyk and Reimann's styles were often not investigated with many alternatives being applied, such as Fiedler's LPC, 'great-person' leader and other styles which were not related to the styles Muczyk and Reimann proposed.

(b) Methodology was often appropriate in many plans who used either participant or non-participant to observe the behaviour of the leader, but sometimes the choice of the observation technique and how it was applied was not explained. The psychological evidence for this question part should have focused specifically on Muczyk and Reimann who proposed that leaders can differ according to 'degree of participation in decision-making' and amount of leader direction/execution of a decision. These two produce four styles: directive autocrat, directive democrat, permissive autocrat and permissive democrat.

Section C

Question 9

A number of very good answers were written, and the debate here often focused on arguments supporting the use of biomedical treatments (drugs) for OCD and arguments against the use of such treatments. A common argument in favour of the use of drugs centred around comorbidity; those with OCD having depression and therefore anti-depressants were useful. However, the same argument was used against the use of drugs because they do not actually treat the OCD. Some answers were weak because a general essay on OCD was written with no attempt to debate or evaluate.

Question 10

There were a few good answers in response to this question, showing that candidates knew about anchoring and heuristics, and the research by Wansink for example, but many answers were weak with candidates not able to demonstrate familiarity with the topic area.

Question 11

This question resulted in many answers describing and then evaluating the life events questionnaire proposed by Holmes and Rahe. This approach scored some bottom-band marks, but this was not the main focus of the question. The debate of the question concerned the use of *psychological* measures of stress, such as the life events questionnaire, being too subjective which is in direct contrast to *physiological* measures of stress, which are objective. Only a few answers addressed this debate and those that did often scored high marks.

Question 12

Answers to this question largely followed the same pattern as seen in the other options, an incorrect structuring of the answer without engagement with the debate. Often there was a description of adaptive leadership. Description can only be credited in ***Section C*** if it is in support of an argument agreeing or disagreeing with the debate presented in the question.

PSYCHOLOGY

Paper 9990/43
Specialist Options: Application

Key messages

- What has been learned from the AS component of the syllabus should be transferred to the A2 component. For example, at AS candidates learn about methodology, such as experiments, which also applies to A2.
- Questions should be read carefully, ensuring that the focus is on what the question asks.
- 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) and a conclusion.
- 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.
- 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.
- Psychological knowledge should be applied wherever possible. Anecdotal and common-sense answers will not achieve top marks.

General comments

There was evidence to suggest that many candidates had not studied two options to the same level. Whilst answers to one option were often very strong, some answers to the second option were notably weaker, often limited to anecdotal or common-sense responses. Further, there were some examples of weak examination technique which candidates would benefit from improving.

Section A

- 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.
- Answers must refer to the study the question is about. Many answers provided general comments which were unrelated to the study itself.
- For **part (d)**, many answers correctly included strengths and weaknesses but often these were not related to the question, and so restricted marks.
- Many conclusions merely repeated what had already been written, and such summaries scored no marks. A conclusion is a 'decision reached by reasoning' and so as the reasoning has been done through the advantages and disadvantages, a final decision/conclusion needs to be drawn.
- Candidates should ensure they are focusing on 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.
- Candidates should provide sufficient detail to score all the available marks. A single sentence is more likely to score one mark rather than two marks, so a little elaboration, explanation or an example that goes beyond the basic sentence is always recommended. Candidates should always try to include psychological knowledge in their responses.

Section B

Answers to **part (a)** questions in this section should include an appropriate design and 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. Describing a relevant piece of research from a topic area alone is not enough to score full marks. The links between the research and how it informed the design must be shown. There is no need to quote a name (date) for each sentence, some responses included 'I chose a self-selecting sample because Milgram (1963) did' for example. This just identifies a study using that technique, rather than explaining 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. Therefore, candidates are not required to describe everything they know about that topic area, and answers that do not address the question will only achieve minimal marks. To score marks at the top end of the mark range, answers must focus on arguments both for and against the statement, and then 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) Many candidates were able to clearly explain the term 'covert sensitisation' and were awarded full marks. A few candidates confused it with imaginal desensitisation, systematic desensitisation and even EMDR. No marks were awarded for incorrect explanation of terms.

(b) (i) Some very strong answers showed good understanding of relevant terms and explained how they were applied in the study by Glover. However, there were also some weak answers which with little knowledge demonstrated about the Glover study. An example of an incorrect response was suggesting that covert means the participant does not know they are being treated.

(c) Many candidates appeared to confuse the terms or did not show understanding of the difference between the terms sensitisation and desensitisation or between imaginal and covert. A few candidates described pictures being shown to a participant, which is incorrect, or defining covert as the participant not knowing that they are involved in treatment, also incorrect. Answers scoring very few marks also described one term followed by the other rather than giving differences as the question required. One difference is that imaginal desensitisation uses progressive muscle relaxation whereas covert sensitisation does not. Another difference is that covert sensitisation creates an unpleasant association (vomiting in the Glover study) whereas imaginal desensitisation reduces the emotions associated with an event.

(d) All **part (d)** questions in ***Section A*** require a discussion of advantages and disadvantages and a conclusion, and this question part was no exception. Further, the advantages and disadvantages must be related to the question, in this instance covert sensitisation. There were a small number of answers achieving full marks, with many candidates not relating their answer to the topic area. Summaries were often included rather than conclusions.

Question 2

(a) The important term that was essential to be addressed in this question was the term 'perceived' because this means that the experience of crowding can vary from one individual to another. Many candidates included this and scored full marks. Others did not show sufficient understanding of the term.

(b) (i) Answers to this question were frequently awarded full marks for stating the difference between the two terms. Candidates understood that human density involves people whereas spatial density is about the physical location, and with regard to consumer psychology, the number of fixtures and fittings of a store, or the amount of merchandise in the store, for example.

(ii) This question asked about the questionnaire used by Machleit et al. to gather data about perceived crowding. Many candidates gave non-specific and general answers such as 'it included a number of questions'. The questionnaire used a Likert-type scale, scale range 1–7, asking questions like 'the store seemed very crowded to me' each of which would score 1 mark.

(c) Many candidates were unable to suggest a model of the effects of ambience and this restricted their marks. Some candidates suggested models of personal space, which also could not be awarded marks. Some candidates suggested the cognition-emotion model, which was applicable, but the strongest answers suggested the Mehrabian and Russell PAD model which is directly applicable. This model considers the levels of Arousal, Pleasure and Dominance (the amount of control a person feels over a situation). This model, applied to the effects of crowding, earned some candidates maximum marks.

(d) Many answers included two strengths and two weaknesses of questionnaires, but only scored partial marks because, frequently, answers were not related to the effects of crowding (or even consumer behaviour) as the question required. An answer might be 'one strength (of using a questionnaire) is that it produces quantitative data which can be statistically analysed'. This is correct but it would not receive full credit without being related to the question, because it could apply to any question from any topic area rather than this specific question.

Question 3

(a) All candidates had some knowledge about the Funhaler device and all scored at least one mark. Answers whose description was more detailed, such as adding more than one point, or that it included a whistle and spinner not included in other devices, were awarded full marks.

(b) At the top end of the mark range, answers were referring to positive reinforcement, Skinner and operant conditioning, and often explained fully how the spinner and whistle were reinforcers. There were some weak answers at the bottom end of the mark range suggesting nothing more than it is a behavioural device because it involves behaviour.

(c) Answers which were not outlined by Watt et al. were acceptable, but candidates suggesting these often failed to provide any more detail than a single sentence. Although this earned one mark (for each suggestion), answers needed detail to elaborate to gain two marks. For example, a one mark answer might be 'children don't take their medicine because they do not like the taste' whereas a two mark answer might be 'children apply rational non-adherence: they do not like the taste of the medicine so make the logical decision that they would rather not take the medicine than suffer the bad taste.'

(d) Many candidates scored no marks or very few marks because they did not answer the question set. The question was specifically 'studies on non-adherence in children' but responses were often a general discussion on using children in psychological studies, using AS level studies rather than those specifically relating to adherence. Any advantage or disadvantage must be focussed on the question in order to achieve marks.

Question 4

(a) There were some excellent explanations of the quality of working life (QWL), most referring to QWL as 'satisfaction with personal and working needs'. Alternatively, there were answers which merely stated that quality of working life is about the quality a person has in their life at work. Repeating or rewording the question cannot be credited.

(b) An outline of two features used to assess QWL were required. Some candidates were not able to demonstrate any knowledge, but others identified a feature (one mark) and then provided an outline to explain the feature (two marks). A maximum mark answer was typically: 'One feature is social integration (one mark) which refers to the interpersonal relationships a worker has with other workers at work' (two marks).

(c) (i) A small number of candidates confused reliability with validity, which couldn't be credited. Some answers were brief, with nothing more written than 'use test-retest' for limited credit. Others wrote 'give the questionnaire to other workers' which is incorrect because to test reliability, the *same* questionnaire should be given to the *same* workers at a later date (i.e. test re-test).

(ii) Many candidates did not make a suggestion and some suggestions were vague, but top marks went to those who suggested using concurrent validity, comparing the results of the QWL test with some other measure of satisfaction. Other measures such as the job descriptive index (JDI) or the Minnesota Satisfaction Questionnaire were appropriate.

(d) This question part followed on from the theme introduced in the stem and asked for advantages and disadvantages of using a five-point scale to measure job satisfaction. Relevant advantages and disadvantages were given, but often these were not always related to the QWL questionnaire (or any other measure). Any advantage or disadvantage must be focussed on the question in order to achieve marks.

Section B

Question 5

(a) This question did not receive many strong responses . The question asked for an investigation into the long-term effectiveness of applied tension. A few candidates knew that applied tension is used to treat blood/injection phobia and planned very good studies. The main weakness was that many candidates did not know what applied tension is used for. It was commonly assumed that it is a treatment for phobias and so a study was planned comparing applied tension with systematic desensitisation (or some other way to reduce anxiety). The problem is that applied tension raises blood pressure whereas other techniques lower blood pressure through relaxation.

(b) The psychological evidence included in this question part needed to focus specifically on applied tension and blood/injection phobia. Whilst a few candidates did this, many candidates wrote about many different kinds of phobias, which was not relevant to the question.

Question 6

(a) Most candidates opted to plan a field experiment, which was a logical choice given the question. Designs were often impressive because the complex IV had to include three conditions of chocolate products placed in three different ways in a film clip. Some opted to place a product at the beginning, in the middle or at the end of a film clip. Others chose to have the product held in the hand, on a table, or an alternative location. The DV was often the number of times the product was noticed when participants were asked in a later questionnaire. Overall, there were some very strong answers.

(b) Methodologically some answers were very strong with detailed explanations for the choice of placement, and many answers also focused on the controls that were applied. Some answers needed to explain the plan in more detail, rather than a single sentence statement. Such answers listed things like 'I used an opportunity sample' without any further comment, but an explanation is required. Psychological knowledge was very strong in many answers, with candidates often explaining how the study on product placement by Auty and Lewis related to their plan.

Question 7

(a) This question on age differences in non-adherence allowed candidates a free choice of method with many choosing to make the study an experiment. However, the consequence of this was that the IV was often muddled between different ages and different reasons. The easier choice, as done by a small number of candidates, was the use of a questionnaire (or an interview), which asked participants of different ages to choose one reason for their non-adherence from a range of different possibilities. Taking time to think about the most appropriate method to use and thinking the plan through before starting to write is time well spent. It is also worth noting that an experiment cannot automatically be applied easily to every question that is set.

(b) Direction to what psychological evidence to use to investigate reasons for non-adherence was provided in the stem of the question, which stated: 'the health belief model provides many reasons

to explain non-adherence'. Yet most candidates overlooked this and invented reasons of their own. These reasons were usually not based on psychological knowledge. A few candidates opted to write about rational non-adherence and this was creditworthy. Methodologically most experiments were not well planned with confused IVs. The methodology for questionnaires was often good, although the reasons for choosing to use open-ended questions for example was often insufficiently detailed.

Question 8

(a) This question also gave a free choice of method, and again an experiment was chosen to be applied by most candidates. Most of these candidates chose to have an IV of 'with targets' and 'without targets' and a DV of 'level of motivation' which was good. What was lacking was knowledge of target setting, such as the five features of SMART targets, with many candidates referring to no more than 'target setting' throughout. The plan should be based on the psychological knowledge explained in **part (b)**. If an experiment is chosen, it is important to include all the specific features (type, IV, DV, controls, design) and also to include a number of general features, such as the sample and sampling technique, the type of data to be gathered and appropriate ethical guidelines.

(b) Methodologically most answers were good but could have been made better with more explanation of the reasons for design decisions. The psychological evidence should have focused on Latham and Locke's goal setting theory and their SMART targets (where targets should be specific, measurable, attainable/agreed, realistic/relevant and time-based). A few answers focused on this work but many did not, instead focusing on the word 'motivation' and writing about Maslow and his hierarchy which was not relevant.

Section C

Question 9

The main debate here was whether the psychoanalytic explanation of phobias could be generalised to everyone, with an additional debate about case studies. There were three types of answers. Firstly, some candidates focused exclusively on the study of Charles, and although advantages and disadvantages of case studies were provided, the psychoanalytic component was absent. Secondly, there was a debate about phobias in general, with reference to little Albert and classical conditioning. Again, there was an absence of the psychoanalytic approach. The third type did include the work of Freud, but often did not go beyond the work of little Hans. Only a few candidates realised that the Psychoanalytic/Freudian explanation of phobias applies to all, suggesting that phobias are due to the id, ego and superego.

Question 10

There were some strong answers in response to this question, showing that candidates knew about eye-tracking, a quantitative, 'scientific' measure that can be used when studying menu item position. The debate was whether this alone was sufficient without the need for qualitative data, or whether people should be asked about what they look at and why when studying a menu. Some detailed answers addressed this debate, but others were superficial, often making vague points that were unrelated to menu design.

Question 11

This question linked unrealistic optimism, outlined by Weinstein with preventing stress. The debate would logically be that those who are unrealistically optimistic would not need to prevent stress, and those that are not unrealistically optimistic would consider doing something to prevent stress. Whilst a few candidates were able to debate this, others struggled due to lack of knowledge about unrealistic optimism, preventing stress or both.

Question 12

The words 'need to achieve' in the question should have focused minds on the syllabus section of achievement motivation and the work of McClelland (1965). McClelland argues that a need for achievement is very important in working life, as is the need for affiliation (to work with and enjoy friendships with work colleagues). Consequently, earning money is less relevant. However, many candidates did not address this and focused on Maslow and his hierarchy of needs.