PSYCHOLOGY

Paper 9990/11
Approaches, Issues and Debates 11

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

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

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

Candidates also need to know the set procedure of studies in the order presented in the original journal article. Questions can be based around just *part* of a procedure and the candidate must be able to produce an answer that is directed and concise rather than writing about the whole of the procedure. In addition, candidates need to know precise details about a procedure. This means presenting a level of detail about the procedure that would mean the study could be replicated. Generic 'stories' about a procedure will not gain credit, candidates still appear not to be so well prepared for these types of question.

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

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

Section B was the strongest predictor of overall success for this examination paper with **Questions 5, 6** and **8** in ***Section A*** also correlating strongly with overall success.

General comments

The marks achieved by the candidates sitting this examination covered a wide spread of possible marks. However, three-quarters of the candidates scored 28 marks or less. Some candidates provided a range of excellent answers to many of the questions and could explain psychological terminology well, providing evidence that they were prepared for the examination.

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

CAMBRIDGE
International Education

© 2024

There were several blank responses in this series. As positive marking is used, candidates should attempt all questions even if they are unsure of the response they are providing.

Many candidates could not differentiate between a result and a conclusion. A result is the collected data from a study that has been analysed via descriptive and inferential statistics. A conclusion is a generic commentary about what the results actually tell us linked to the aim/purpose of the study. This has been highlighted in previous PERTs but many candidates get these types of question incorrect.

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

Comments on specific questions

Question 1

(a) The minority of responses could correctly state the number of monkeys used in the final sample. Common errors included stating the number of monkeys used in other parts of the study, for example, 82.

(b) Stronger responses could clearly identify two 'wheeled' toys. Popular choices included wagon, car and truck. Common errors included naming 'plush' toys or naming specific cars and trucks that were not used in the study. It is essential for candidates to know how materials were used in any core study.

(c) A large majority of responses presented a result rather than a conclusion. Common results were linked to toy choice/frequency/preference. These could not be awarded any credit. A small minority of responses could provide the generic conclusion based on the results presented by Hassett et al. Conclusions need to be based around psychological principles and assumptions that are logical and meaningful. Knowing that male monkeys preferred wheeled toys the most may be useful in other domains, but not for the study by Hassett et al. Candidates need to know the difference between results and a conclusion. A result is the collected data from a study that has been analysed via descriptive and inferential statistics. A conclusion is a generic commentary about what the results actually tell us linked to the aim/purpose of the study.

Question 2

(a) A minority of responses could provide a full result based on the change of distress ratings for the given scenario. The most popular choice was focused on the continual reduction of distress scores over time. A significant minority of responses stated that the ratings went up which is incorrect as this was for a different scenario. It is very important that candidates know all main results from all core studies

(b) The majority of responses could identify a weakness of the study by Saavedra and Silverman. Popular choices included lack of generalisability and ethical issues. There was a substantial minority of responses that only gave a generic weakness without any context. The question ended with '…of this study', therefore candidates need to be aware that one mark is for identifying an appropriate weakness and one mark is for evidence from the study to show the Examiner why it is a weakness.

Question 3

(a) A minority of responses could provide a meaningful aim of the study by Pozzulo et al. The most common choice was that they were investigating the role of both cognitive and social factors linked to false positives given by children. Many responses focused on the testing of negative/positive false responses to human and cartoon faces. This was how the aim was tested (predictions) and not the aim of the study. The false responses are predictions based on the generic aim of cognitive/social factors affecting choices in a line-up. It is essential that candidates know the difference between an aim and a hypothesis/prediction where appropriate.

(b) A minority of responses could clearly explain why the study by Pozzulo et al. is from the cognitive approach. Stronger responses presented a key finding and then explicitly linked it to an assumption

© 2024

of the cognitive approach. Examples of these are on the mark scheme. However, there were many responses that gave generic accounts of the study and how it was about 'memory' which could only gain partial credit and the command word was 'explain' so the responses had to cover 'why' the study by Pozzulo et al. is from the cognitive approach. It is essential for candidates to be able to have examples of why each core study has been placed in one of the four approaches on the syllabus.

Question 4

Stronger responses here were focused on the learning task that was used in the study by Milgram. Popular choices included the use of word pairs, how the learner had to indicate the second word via a switch with a quadrant lighting up for the teacher to see. However, the majority of responses focused on the shock instructions and not the learning task as covered in the original paper by Milgram. It is essential for candidates to know 'each piece' of the procedures for all 12 core studies rather than just knowing a generic 'story' about what participants were expected to do in each of them. Questions can be asked about any part of the procedure.

Question 5

(a) Stronger responses could clearly outline two different assumptions of the biological approach. Popular choices focused on the role of the brain/genetics/hormones/evolution on human behaviour. These assumptions are in the syllabus so it is essential that candidates know them and can present them coherently to an Examiner. Common errors included presenting assumptions from a different approach or writing the same assumption in a different way – in these cases only the highest scoring assumption would be credited.

(b) A minority of responses were credited with maximum marks here by presenting an appropriate finding (a common one was that grey matter increased in the MBSR group) and then linking it back explicitly to one of the assumptions they had presented in **Question 5a**. A significant proportion of candidates could provide a correct finding but then simply wrote out the assumption from **Question 5a** that could explain the findings. However, this is implicit so the second mark could not be awarded. The second mark could only be awarded when a candidate explicitly explained why the assumption had been supported. This had the highest rate of blank responses on this paper.

Question 6

(a) A minority of responses could clearly describe how a participant received the oxytocin; via nasal drops. However, there were some alarming errors here where candidates were stating that participants were injected, took pills, were strapped to a chair and given it, or being made to breathe it. Candidates need to know how trials like this are ethically run in psychological studies. This question has the third highest rate of blank responses.

(b) The majority of candidates could correctly identify that the other condition was a placebo one. However, there were some alarming errors here including dopamine, serotonin, alcohol and other drugs. This had the second highest rate of blank responses.

(c) Similar to **Question 2b**, many responses could identify an appropriate strength of the study by Perry et al. Popular choices included standardisation/reliability, use of a double-blind technique, and explicit ethical strengths. However, like with **Question 2b**, only a minority of responses explained the strength via an explicit example from the study. Candidates need to be aware that when a question has '…..this study' within it, an explicit example from the study is needed to be able to be awarded maximum credit.

Question 7

The majority of responses provided some examples of what Gabriela could do in her job based on the presented scenario. However, many responses presented an implicit 'how' in terms of stating 'allow doodling in the job'. The term 'doodling' was in the question, so candidates needed to go beyond copying a word to show explicit understanding of what is involved in doodling to be awarded any credit. Stronger responses provided explicit information about how to doodle including providing paper materials, having shapes to shade etc. This showed explicit knowledge of what is involved in doodling so could be awarded up to four marks. It is vitally important for candidates to present explicit information to Examiners to clearly show understanding of concepts already given in a question.

© 2024

Question 8

There were a range of responses to this question. Stronger responses could clearly argue why Adamu had stated that the study was not generalisable. The more popular choices included focusing on the demographics of the AS/HFA group, but also Group 3. However, the majority of responses only described lack of generalisability, rather than explaining. For example, a candidate might correctly describe that Group 1 was all male and stop there. However, this is not answering the question as to 'explain' means to tell the Examiner who it then might be difficult to generalise to, in this case, females with AS/HFA. Implicit statements cannot be awarded credit as it does not show the Examiner any depth of understanding about generalisability. For questions like this, candidates need to be able to provide correct examples and arguments and be explicit in their explanations.

Question 9

(a) A significant minority of responses could outline one 'vertical' eye movement dream reported in the study by Dement and Kleitman with a smaller proportion outlining the required two dreams. Popular correct choices included climbing a ladder and playing basketball. Many responses mixed together the reported dreams, for example, climbing a ladder at the bottom of a cliff. Some candidates outlined 'horizontal' eye movement dreams which could not be credited. Some candidates created dreams that were not reported in the study by Dement and Kleitman.

(b) Stronger responses could clearly explain two differences, including the named comparison. Popular choices included the compulsory brain measurement techniques, ethics and type of data collected. To improve responses to this type of question, candidates need to choose comparison points that can be developed and explained, using examples from both studies to explain each difference. For example, explaining the use of different brain measurement techniques with a focus on the function of each technique with clear examples from both studies. However, stating that each study had a different aim does not allow the response to be detailed so will always only achieve Level 1. Candidates need to choose carefully what the comparisons are, ensuring that they are logical and can be explained fully, using examples from both studies. It is also very important to read the question to see what can or cannot be used on the response. In this case, the candidates were told to refer to the brain measurement techniques, yet a minority of candidates did not use the brain measurement techniques in their responses and were therefore awarded the Level of their best 'difference' only. There was also a sizeable minority of responses that could not state the correct brain measurement technique used in one or both of the named studies.

Question 10

The strongest responses evaluated the study by Bandura et al. in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of validity. Common choices included ethics, generalisability, observations, reliability, and quantitative data. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Bandura et al. to explicitly support their point. These answers tended to score Level 5 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses, all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by 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 not in context which meant it could not be given Level 5. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills *only*. In addition, some responses are still following a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). A response that fails to have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did use validity data in an evaluative sense but did not fully explain why it could be a strength and/or a weakness or simply stated 'therefore it lacked validity'. Some responses did not cover the named issue. There were also a large number of factual errors about the study by Bandura et al. presented as facts by candidates including the children watched videos, the children gave valid consent and parents gave valid consent. It is essential that candidates choose evaluation points based on what actually happened in the study. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, avoiding real world application. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples (plural) from the study to show clear

© 2024

understanding. An evaluation that is in depth tends to have at least two explicit examples from the named study for every evaluative point made. These are the requirements for a Level 5 response. The average response was Level 2 for this cohort.

4Uadmission
SINCE 2020

© 2024

PSYCHOLOGY

Paper 9990/12
Approaches, Issues and Debates 12

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 relating back to a previous answer. To achieve full marks, these need to be correctly present in their responses. The essay (final question) requires four evaluation points to be in depth (two strengths and two weaknesses) with at least one of these about the named issue. In depth tends to mean having two examples of a particular concept to support an evaluative point. Credit is limited if the named issue is omitted or just described. In addition, candidates need to ensure that the concept they are using for an evaluative point is one that is relevant and applicable.

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

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

Candidates also need to know the set procedure of studies in the order presented in the original journal article. Questions can be based around just *part* of a procedure and the candidate must be able to produce an answer that is directed and concise rather than writing about the whole of the procedure. In addition, candidates need to know precise details about a procedure. This means presenting a level of detail about the procedure that would mean the study could be replicated. Generic 'stories' about a procedure will not gain credit, candidates still appear not to be so well prepared for these types of question.

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

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

Section B was the strongest predictor of overall success for this examination paper with **Questions 2**, **5** and **6** in ***Section A*** also correlating strongly with overall success.

General comments

The marks achieved by the candidates sitting this examination covered a wide spread of possible marks. However, half of the candidates scored 24 marks or less. Some candidates provided a range of excellent answers to many of the questions and could explain psychological terminology well, providing evidence that they were prepared for the examination.

Stronger overall responses followed the demands of each question with explicit use of psychological terminology and logical, well-planned answers in evidence. Appropriate examples were used from studies

© 2024

when the question expected it and there was evidence of candidates being able to apply their knowledge to real-world behaviours in terms of what and how.

There were several blank responses in this series. As positive marking is used, candidates should attempt all questions even if they are unsure of the response they are providing.

Many candidates could not differentiate between a result and a conclusion. A result is the collected data from a study that has been analysed via descriptive and inferential statistics. A conclusion is a generic commentary about what the results actually tell us linked to the aim/purpose of the study. This has been highlighted in previous PERTs but many candidates get these types of question incorrect.

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

Comments on specific questions

Question 1

(a) A large majority of responses could correctly state that the boy was being treated for button phobia. Common errors included naming a different phobia.

(b) Stronger responses could clearly describe components of the Feelings Thermometer. Popular choices were it being a 9-point scale and that it was a hierarchy produced by the boy. Common errors included stating the incorrect range used on the thermometer or giving a result from the treatment sessions.

(c) A minority of responses could identify two correct findings. Popular choices included no longer meeting DSM criteria and that he could now wear plastic buttons. Common errors included giving a generic idea about how the boy was now behaving giving findings unique to the other follow-up sessions. It is essential for candidates to know all key results from all 12 core studies.

Question 2

(a) A minority of responses could provide a full result based on the sex of the participant on the eyes test. The most popular choice was females scoring higher than males. However, this could only be awarded maximum credit if the groups were correctly identified which happened infrequently. As with previous series', there can be no credit given to a value judgement, in this case, one group being better or worse. A result is the reporting of factual statistical information with no interpretation. It is very important that candidates know to present results with meaningful comparisons with no interpretation.

(b) The majority of responses could identify a weakness of the study by Baron-Cohen et al. Popular choices included lack of generalisability, lack of mundane realism, and ethical issues. There was a substantial minority of responses that only gave a generic weakness without any context. The question ended with '…of this study', therefore candidates need to be aware that one mark is for identifying an appropriate weakness and one mark is for evidence from the study to show the Examiner why it is a weakness.

Question 3

(a) A slight minority of responses could provide a partial definition of the term bystander apathy. There are two components of the definition, one for each word, and the most popular correct part was for apathy being less likely to help out. However, many responses did not define the term bystander or confused apathy with empathy. It is important that candidates know the definitions of key terms from the psychology being investigated section of the syllabus for each of the 12 core studies.

(b) A minority of responses could clearly explain one reason why the study by Piliavin et al. could support a situational explanation. Stronger responses presented a key finding and then explicitly linked it to why it shows a situational explanation. Examples of these are on the mark scheme. However, there were many responses that gave a tautological explanation by stating 'it supports the situational explanation of behaviour as the situation affected behaviour'. Response like this example cannot be given credit as it is not certain to an Examiner whether the candidate clearly

© 2024

understands the term 'situation'. It is essential that candidate do not provide tautological definitions to terms.

Question 4

Stronger responses here were focused on the inclusion criteria used in the study by Hölzel et al. Popular choices included having no issues with an MRI, having not participated in meditation for at least six months, and that they were healthy either psychologically or physical. However, the majority of responses focused on the procedure for the entire study or giving a detailed account of what was in the MBSR course. It is essential for candidates to know 'each piece' of the procedures for all 12 core studies rather than just knowing a generic 'story' about what participants were expected to do in each of them. Questions can be asked about any part of the procedure. This question had the highest rate of blank responses.

Question 5

(a) Stronger responses could clearly outline two different assumptions of the social approach. Popular choices focused on how other people can influence our behaviours and emotions. These assumptions are in the syllabus so it is essential that candidates know them and can present them coherently to an Examiner. Common errors included presenting assumptions from a different approach or giving tautological assumptions that behaviour is affected by social factors. For the latter these need to be explicitly identified to be awarded credit.

(b) A minority of responses were credited with maximum marks here by presenting an appropriate finding (a common one focused on the different personal space levels for the imagined 'people' approaching) and then linking it back explicitly to one of the assumptions they had presented in **Question 5a**. A significant proportion of candidates could provide a correct finding but then simply wrote out the assumption from **Question 5a** that could explain the findings. However, this is implicit so the second mark could not be awarded. The second mark could only be awarded when a candidate explicitly explained why the assumption had been supported. This had the second highest rate of blank responses on this paper.

Question 6

(a) A slight minority of responses could clearly describe the procedure of a trial used in the study by Hassett et al. However, there were many responses that focused on what happened after the trial had ended, for example, how the behaviour was coded or timed. It is essential for candidates to read the question carefully and reflect on what the question is asking the response to be focused on before starting to answer. Stronger responses clearly understood that the focus needed to be on what happened during a trial rather than after a trial had ended.

(b) Similar to **Question 2b**, many responses could identify an appropriate strength of the study by Hassett et al. Popular choices included focusing on it being covert and recorded. However, like with **Question 2b**, only a minority of responses explained the strength via an explicit example from the study. Candidates need to be aware that when a question has '…..this study' within it, an explicit example from the study is needed to be able to be awarded maximum credit. Also, a minority of responses presented a correct strength of the study, but it was not linked to observations as required by the question. It is important for candidates to read the question carefully and focus their response accordingly.

Question 7

The majority of responses provided some examples of what could be used with children based on the presented scenario. Popular choices included making sure the police were dressed 'casually', never having a target absent line-up, and do not question a child's decision so as to eliminate any doubts. However, some responses gave generic accounts with no explicit link to the study by Pozzulo et al. In addition, many responses simply provided an account of the study by Pozzulo et al. with no link to the scenario in the question. It is vitally important for candidates to present explicit information to Examiners to clearly show understanding of concepts already given in a question.

Question 8

There were a range of responses to this question. Stronger responses could clearly argue why Tejas had stated that the study was not ethical. The more popular choices included focusing on psychological harm, not

© 2024

gaining valid consent, and deception. Stronger responses could clearly identify a range of ethical issues and provide explicit examples from the study by Bandura et al. For example, the children did not given informed consent, they were exposed to adults behaving aggressively, and the children could have continued with aggression post-study. For questions like this, candidates need to be able to provide correct examples and arguments and be explicit in their explanations. There were some worrying misconceptions presented in some answers in terms of candidates claiming parents gave valid consent (not answering the question anyway) and that the children watched videos of adults. It is essential that candidates are taught the correct study by Bandura et al. as explicitly stated in the syllabus.

Question 9

(a) A significant majority of responses could outline at least three features of the sample used in the study by Milgram. 'Popular choices included the sample size, that they were all male, that they were recruited via volunteer sampling, and their age range. A minority of responses described the sample from a different core study or claimed that Milgram used males and females. It is essential for candidates to know the characteristics and features of the sample for all 12 core studies.

(b) Stronger responses could clearly explain two differences, including the named comparison. Popular choices included the compulsory ethics, and the research method used. To improve responses to this type of question, candidates need to choose comparison points that can be developed and explained, using examples from both studies to explain each difference. For example, explaining that one study was more/less ethical than the other with clear examples from both studies. However, stating that each study had a different aim does not allow the response to be detailed so will always only achieve Level 1. Candidates need to choose carefully what the comparisons are ensuring that they are logical and can be explained fully, using examples from both studies. It is also very important to read the question to see what can or cannot be used on the response. In this case, the candidates were told to refer to ethics, yet a minority of candidates did not use ethics in their responses and were therefore awarded the Level of their best 'difference' only. There was also a sizeable minority of responses that attempted to compare the studies on similarities (e.g. sampling technique). Candidates need to be careful to read whether it is two differences, two similarities, or one difference and one similarity that is required.

Question 10

The strongest responses evaluated the study by Andrade in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of the sampling technique used. Common choices included generalisability, validity, reliability, and quantitative data. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Andrade to explicitly support their point. These answers tended to score Level 5 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses, all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by Andrade as examples which meant the response scored in the lower bands. Other responses included three evaluation points that were thorough, logical, and well argued with a fourth point that was not in context which meant it could not be given Level 5. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills *only*. In addition, some responses are still following a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). A response that fails to have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did use the sampling technique in an evaluative sense but did not fully explain why it could be a strength and/or a weakness or presented an incorrect sampling technique. Some responses did not cover the named issue. There were also a large number of conceptual errors about the study by Andrade presented as facts by candidates including the surprise test being psychologically damaging (there was a large ceiling effect to negate this), that all participants were candidates, or that demand characteristics affected results (on p.103 of the Andrade study it clearly states that those who suspected a memory test, when their data were removed, it did not alter the pattern of results). It is essential that candidates choose evaluation points based on what actually happened in the study, and/or what is applicable. Not all psychological concepts are relevant to every one of the 12 core studies. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, avoiding real world application. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples (plural) from the study to show clear understanding. An evaluation that is in depth tends to have at

© 2024

least two explicit examples from the named study for every evaluative point made. These are the requirements for a Level 5 response. The average response was Level 2 for this cohort.

4Uadmission

© 2024

PSYCHOLOGY

Paper 9990/13
Approaches, Issues and Debates 13

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 relating back to a previous answer. To achieve full marks, these need to be correctly present in their responses. The essay (final question) requires four evaluation points to be in depth (two strengths and two weaknesses) with at least one of these about the named issue. In depth tends to mean having two examples of a particular concept to support an evaluative point. Credit is limited if the named issue is omitted or just described. In addition, candidates need to ensure that the concept they are using for an evaluative point is one that is relevant and applicable.

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

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

Candidates also need to know the set procedure of studies in the order presented in the original journal article. Questions can be based around just *part* of a procedure and the candidate must be able to produce an answer that is directed and concise rather than writing about the whole of the procedure. In addition, candidates need to know precise details about a procedure. This means presenting a level of detail about the procedure that would mean the study could be replicated. Generic 'stories' about a procedure will not gain credit, candidates still appear not to be so well prepared for these types of question.

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

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

Section B was the strongest predictor of overall success for this examination paper with **Questions 4**, **5** and **6** in ***Section A*** also correlating strongly with overall success.

General comments

The marks achieved by the candidates sitting this examination covered a wide spread of possible marks with half of the candidates scoring 29 marks or more. 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

© 2024

when the question expected it and there was evidence of candidates being able to apply their knowledge to real-world behaviours in terms of what and how.

There was a very low level of blank responses. As positive marking is used, candidates should attempt all questions even if they are unsure of the response they are providing.

Many candidates could not differentiate between a result and a conclusion. A result is the collected data from a study that has been analysed via descriptive and inferential statistics. A conclusion is a generic commentary about what the results actually tell us linked to the aim/purpose of the study. This has been highlighted in previous PERTs but many candidates get these types of question incorrect.

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

Comments on specific questions

Question 1

(a) A majority of responses could correctly state the number of participants who had minimal data collected from them. Common errors included repeating the number given in the **Question (5)** or providing a sample size from a different core study.

(b) Stronger responses could clearly state the other two instructions given to participants. Common errors included stating something that happened after the participants had begun the study. It is essential that candidates read the question carefully to ensure they are focusing on the correct part of the study. This had the second highest rate of blank responses on this paper.

(c) The average score for this question was 1 mark. This meant that candidates tended to provide partial or brief conclusions. Popular two-mark answers were the eye movements corresponding to the content of the dream and that dreams are experienced in real-time. Common errors included presenting results instead of a conclusion. Candidates need to know the difference between results and a conclusion. A result is the collected data from a study that has been analysed via descriptive and inferential statistics. A conclusion is a generic commentary about what the results actually tell us linked to the aim/purpose of the study.

Question 2

(a) A large minority of responses could provide a full result based on the sex differences and plush toys. The most popular choice was females playing with plush toys more often. Common errors were based around not reading the question carefully. For example, responses focusing on wheeled toys and comparing female play with wheeled and plush toys. These are not answering the question set. It is essential for candidates to read the question carefully to ensure they are presenting the correct results. A result is the reporting of factual statistical information with no interpretation. It is very important that candidates know to present results with meaningful comparisons with no interpretation.

(b) The minority of responses could identify a weakness of the study by Hassett et al. and provide an explicit example. Popular choices included lack of generalisability and lack of mundane realism. There was a substantial minority of responses that only gave a generic weakness without any context. The question ended with '…of this study', therefore candidates need to be aware that one mark is for identifying an appropriate weakness and one mark is for evidence from the study to show the Examiner why it is a weakness.

Question 3

(a) A slight majority of responses could provide a correct outline of how positive reinforcement was used. There are two components linked to this of who provided it and when. The latter was the most popular feature of responses. However, many responses did not note that the mother was the only one to provide this type of reinforcement.

© 2024

(b) A minority of responses could clearly explain one reason why the study by Saavedra and Silverman could support the nurture side of the nature versus nurture debate. Stronger responses presented a key finding and then explicitly linked it to why it shows a nurture explanation. Examples of these are on the mark scheme. However, there were many responses that gave a tautological explanation by stating 'it supports the nurture explanation of behaviour as the behaviour was nurtured'. Responses like this example cannot be given credit as it is not certain to an Examiner whether the candidate clearly understands the term 'nurture'. It is essential that candidates do not provide tautological definitions to terms.

Question 4

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

Question 5

(a) Stronger responses could clearly outline two different assumptions of the learning approach. Popular choices focused on the role of social learning or operant conditioning in human behaviour. These assumptions are in the syllabus so it is essential that candidates know them and can present them coherently to an Examiner. Common errors included presenting assumptions from a different approach or giving tautological assumptions that behaviour is learnt. For the latter these 'learning mechanisms' need to be explicitly identified to be awarded credit.

(b) A slight majority of responses were credited with maximum marks here by presenting an appropriate finding (a common one focused on children being more likely to show aggression after witnessing an aggressive model than a non-aggressive model) and then linking it back explicitly to one of the assumptions they had presented in **Question 5a** (in this case social learning theory). A significant proportion of candidates could provide a correct finding but then simply wrote out the assumption from **Question 5a** that could explain the findings. However, this is implicit so the second mark could not be awarded. The second mark could only be awarded when a candidate explicitly explained why the assumption had been supported.

Question 6

(a) A minority of responses could describe more than one aspect of the psychology being investigated in the study by Andrade. Popular choices included attention, arousal and information processing. However, a large majority of responses focused entirely on what Andrade did in their study, rather than providing a generic account of the psychology being investigated. To improve, candidates need to know the psychology being investigated listed beneath all 12 core studies in the syllabus, generically. There is only one mark available in these types of questions for an explicit example from the named core study.

(b) Similarly to **Question 2b**, many responses could identify an appropriate strength of the study by Andrade in relation to reliability. Popular choices included focusing on aspects of the study that were standardised. However, like with **Question 2b**, only a minority of responses explained the strength via an explicit example from the study. Candidates need to be aware that when a question has '… this study' within it, an explicit example from the study is needed to be able to be awarded maximum credit. Also, a minority of responses presented a correct strength of the study, but it was not linked to reliability as required by the question. It is important for candidates to read the question carefully and focus their response accordingly. Finally, standardisation by itself is descriptive so cannot gain credit in evaluation-based questions. The consequences of standardisation which are replication to test (not improve) reliability is evaluative.

© 2024

Question 7

There were a range of responses to this question. Stronger responses could clearly argue to Ruru why the nurse had given incorrect medicine. The more popular choices included the doctor being an authoritative figure and that he may have been stern wearing a uniform. Stronger responses could clearly identify a range of reasons based on examples from the study by Milgram. For questions like this, candidates need to be able to provide correct examples and arguments and be explicit in their explanations.

Question 8

There were a range of responses to this question. Stronger responses could clearly argue why Betsy had stated that the study had validity. The more popular choices included focusing on the objectivity of brain scans and that the MBSR program was in a real-life setting for that group of participants. Stronger responses could clearly identify a range of 'validity' points and provide explicit examples from the study by Hölzel et al. For questions like this, candidates need to be able to provide correct examples and arguments and be explicit in their explanations.

Question 9

(a) A significant majority of responses could outline <u>at least three features</u> of the sample used in the study by Fagen et al. Popular choices included the sample size, the age distribution, that they were from the same stable and that there were selected as they were docile. A minority of responses described the sample from a different core study or claimed that Fagen used a much higher sample size of elephants or a different species. It is essential for candidates to know the characteristics and features of the sample for all 12 core studies.

(b) Stronger responses could clearly explain one similarity and one difference. Popular choices included ethics, type of data collected and mundane realism. To improve responses to this type of question, candidates need to choose comparison points that can be developed and <u>explained</u>, using examples from both studies to explain each similarity or difference. For example, explaining that one study was more/less ethical than the other with clear examples from <u>both</u> studies. However, stating that each study had a different aim does not allow the response to be detailed so will always only achieve Level 1. Candidates need to choose carefully what the comparisons are ensuring that they are logical and can be explained fully, using examples from both studies. It is also very important to read the question to see what can or cannot be used on the response. In this case, the candidates were told not to refer to the sample, yet a minority of candidates did use the sample in their responses. There were some alarming misconceptions presented by candidates including that Fagen was a field experiment and that Fagen used an independent groups design.

Question 10

The strongest responses evaluated the study by Piliavin et al. in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of ethics. Common choices included ethics, generalisability, validity, reliability, and quantitative data. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Piliavin et al. to explicitly support their point. These answers tended to score Level 5 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses, all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by 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 not in context which meant it could not be given Level 5. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills *only*. In addition, some responses are still following a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics).

A response that fails to have one evaluation point about the named issue can only score Level 3 (6 marks) maximum. There were many responses that briefly outlined strengths and weaknesses with only some being in context which is a Level 2 response. Any response that has no context cannot get above a Level 1 mark. In addition, many responses did ethics in an evaluative sense but did not fully explain <u>why</u> it could be a strength and/or a weakness or presented incorrect information. Some responses did not cover the named issue. There was one misconception that appeared in some candidates' responses. Some were claiming that inter-observer reliability was excellent as there were two observers. Whilst it is correct that two observers were used, they were coding <u>different</u> behaviours/measures so there was no attempt at testing inter-observer

© 2024

reliability. It is essential that candidates choose evaluation points based on what actually happened in the study, and/or what is applicable. Not all psychological concepts are relevant to every one of the 12 core studies.

To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue, avoiding real world application. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with examples (plural) from the study to show clear understanding. An evaluation that is in depth tends to have at least two explicit examples from the named study for every evaluative point made. These are the requirements for a Level 5 response. The average response was Level 3 for this cohort.

4Uadmission
SINCE 2020

© 2024

PSYCHOLOGY

Paper 9990/21
Research Methods 21

Key messages

- Candidates need to ensure that their responses are focused on the questions within the exam paper. There was more than one instance where it was clear that candidates had misread the question and provided responses which were not creditworthy.
- Candidates need to ensure that they understand the expectations for different command words used on the paper; for example, describe and explain. For describe, candidates need to ensure that they provide a sufficient number of unique points related to the marks allocated to the question, whereas for explain candidates need to identify a particular feature/concept/theory and then **link** their detail point to the feature/concept/theory they have identified. Often, explain questions had poorer outcomes in terms of marks as a candidate just described.
- Candidates need to ensure that they are able to define/outline key terms within the syllabus such as 'matched pairs design' and know the difference between terms such as inter-rater/inter-observer reliability.
- Candidates need to ensure that they link their answers to the information given in the stem if asked to do so. Often candidates showed excellent understanding of named issues/studies but then lost marks for giving generic responses.
- It is worth noting that candidate responses for the 10-mark extended response question showed good knowledge and understanding of questionnaires. There were a number of thoughtful responses for this extended response question, and candidates should be commended for their performance.
- Candidate performance on the 6-mark question about inter-rater and inter-observer reliability was below expectation. It was clear that many candidates did not understand the difference between the two and struggled to coherently explain what the terms mean. This is often a fairly difficult concept to grasp for candidates and therefore some time needs to be taken to ensure that candidates understand these very important terms and can show this understanding in written form.

General comments

This was the first November series for the new Psychology syllabus. Paper 21 scripts provided the full range of marks, showing a good level of knowledge and understanding across many areas of the specification. Where performance was limited, it was due to a lack of knowledge of key terms, or a misunderstanding of the demands of the question. This was clear when looking at questions on matched pairs design (**1**), inter-rater/observer reliability (**5**), and situational variables (**7bi** and **7bii**). Candidate responses showed good knowledge and understanding of key studies such as when referring to the Baron-Cohen et al. (aggression) study, and the Hassett et al. (monkey toy preferences) study, although responses did show some misunderstanding in relation to the independent variables of the Hassett at al. study. Candidates did show their ability to structure 10-mark responses **(9a)** with many able to produce thoughtful procedures which incorporated all the elements required within the question. It is clear that centres have prepared candidates well for the exam. For future series, candidates need to ensure that they have a good understanding of command words, key research method terminology, such as validity and reliability, and the studies which have been named on the syllabus.

CAMBRIDGE
International Education

© 2024

Comments on specific questions

Section A

Question 1

This two-mark question required candidates to explain what is meant by a 'matched pairs design', including any example. To achieve two marks candidates needed to show their knowledge of matched pairs design, and then give an appropriate example. A full range of marks was seen for this question, with many candidates showing good knowledge and understanding.

Where candidate responses did not achieve the two marks it was due to a misunderstanding of matched, or vague responses which did not make it clear that the participants with similar characteristics were placed into different levels of the independent variable. Often candidates would use the words 'matched' participants, rather than participants with similar characteristics, which does not really show a level of knowledge and understanding which is creditworthy.

The other common error is that candidates would show their understanding of matched pairs but then forgot to give a relevant example; meaning they could only achieve one mark.

Question 2

This two-mark question asked candidates to identify the type of hypothesis in this experiment and justify their answer. To achieve the two marks candidate responses needed to identify the correct hypothesis (directional/alternative) and then give a justification of why this is the case. Most candidate responses gave the correct type of hypothesis however, they were then unable to give an appropriate justification meaning that one mark was the most common outcome for this question.

Where candidates gave alternative hypothesis as their response, they more often gave a relevant justification. Candidates found justifying why it is a directional hypothesis more difficult, with many just repeating the hypothesis itself which was not creditworthy. Candidates often find these types of questions more of a challenge and therefore teachers could use past papers and unseen scenarios to allow candidates to practise this type of response

Question 3

(a) This one-mark question asked candidate to define what is meant by the term 'ecological validity'. There were no major issues with this question, with most candidates able to show accurate understanding of the term. Where candidates did not achieve the one mark available it was due to concentrating on the situation being natural to the participants which is not creditworthy as an explanation of ecological validity.

(b) (i) This one-mark question asked candidates to outline one problem with the ecological validity of the revised eyes test. To achieve the mark available candidates needed to outline such problems as; only seeing the eyes, that the eyes were static, that body language could usually be observed. Most candidates were able to achieve the mark available on this question, with the most common response being the issue of the eyes in the test being static. Where the mark was not achieved it was due to candidates concentrating on where the test was taken, rather than the eye test itself, which was not creditworthy.

(ii) This one-mark question asked candidates to explain one way in which the problem you identified in **part bi** could be solved. Performance on this question was directly related to how candidates responded to **part 3bi**. Candidates who suggested static eyes, or only eyes shown, inevitably achieved the mark available on this question. Where candidates did not achieve the mark, it was due to their responses on **part 3bi** concentrating on where the test was taken, therefore the solving of this issue was not creditworthy either.

Question 4

(a) This two-mark question asked candidates to describe what is meant by an 'experiment'. To achieve two marks candidates needed to describe two out of the three definitive points which are on the syllabus related to experiments. Most candidates were able to achieve at least one mark for this question, with the use of a manipulated independent variable and a measured dependant variable

© 2024

being the most common response. Unfortunately, many candidates were then unable to describe a second unique point limiting them to the single mark. The most common error in responses was talking about the use of a hypothesis which is not creditworthy and therefore meant candidates were only able to achieve the one mark.

(b) This one-mark questions asked candidates to identify an independent variable from the Hassett et al. (monkey toy preferences) study. There were no major issues with this question with most candidates able to identify an accurate independent variable from the study. The most common responses were the type of toy and the gender of the monkeys. Where candidates did not achieve the mark, it was due to confusion between the independent variable and the dependent variable with some suggesting that the IV was the toy picked by the monkeys, which was not creditworthy.

(c) This four-mark question asked candidates to describe to strengths of experiments as a research method. To achieve four marks candidates had to identify two strengths of the experimental method and then give some further detail about the strength identified. The full range of marks was seen for this question.

The majority of candidates were able to achieve some marks on this question. The most common strengths identified were the ability to standardise the procedure, and the ability to control extraneous variables so that it is only the IV that affects the DV. Where these strengths were identified most candidates were able to go on and give some detail such as the ability to replicate the study to test reliability and would often achieve full marks.

The most common error was candidates using terms alone i.e. reliability and validity without giving any further detail. Candidates could achieve marks for terms but ONLY when used as a point of detail after a relevant strength was identified as shown above. Candidates need to ensure they do not use terms alone as this does not show any understanding at all in terms of methodological strengths and weaknesses

Question 5

This six-mark question required candidates to describe inter-rater reliability and inter-observer reliability, using any example(s). Awarded marks could come from accurate descriptions of each term and appropriate examples such as the use of two observers in the Fagen et al. study. Candidates could also achieve marks by stating how reliability could be improved. Candidates found this question quite challenging, with many struggling to differentiate between inter-rater and inter-observer reliability.

At the lower end of the mark range, candidates usually achieved one or two marks for suggesting inter-rater reliability was often used with questionnaires or tests. They may be able to give an appropriate example such as from the Bandura et al. study where they found aggression ratings similar between researchers. At the lower mark range differentiation between the two types of reliability was often not present or confused.

At the high end of the mark range candidates produced some thoughtful responses, were able to differentiate between inter-rater and inter-observer reliability and use relevant examples to further show their understanding of the terms. At this range, candidates would also suggest ways of ensuring high reliability such as operationalisation of categories. It is worth remembering that examples do not have to be from named studies, and there were some creative responses which used a variety of examples, which are indeed creditworthy.

Lastly, it is also worth noting that responses cannot be credited twice for the same information no matter if it is used in different techniques; such as suggesting that you would obtain a strong positive correlation if results from different researchers were similar.

Question 6

(a) (i) This one- mark question asked candidates to suggest one pleasant scene that Dr Eynon could use for one of her photographs, other than a beach. There were no issues with this question and almost all candidates were able to achieve the one mark available. Where candidates did not achieve the mark, it was due to suggesting a beach which was not creditworthy as per the question or suggesting something which was clearly not pleasant. This was extremely rare, however.

(ii) This two-mark question asked candidates to suggest two features that Dr Eynon should control about the scenes. To achieve the two marks available candidates needed to suggest two different

© 2024

features which would ensure similarity between the scenes. Candidates performed well on this question, with some really thoughtful responses. The most common answers were suggestions about time of day, weather and colour. There were some creative suggestions about the angle that the picture was taken at which were creditworthy.

Where candidates did not achieve the two marks it was due to responses suggesting the number of people in the pictures. As one of the scenes did not feature people this could not be credited, however if they specified that there was a need to control the number of people in the scenes WITH people in then this could be credited.

(iii) This two-mark question asked candidates to explain why one of the features suggested in **part aii** would be important in Dr Eynon's study. To achieve the two marks candidates needed to choose one of the two features identified in **part 6aii** and say why this was important and then give further detail to exemplify on this point. Candidate performance on this question was directly related to their success on **part ii**. If candidates had achieved no marks on **part 6aii** then inevitably they would achieve no marks on this part of the question.

If candidates had achieved one mark on **part 6aii**, then their ability to achieve full marks on this part depended on which feature they chose. Invariably, where candidates achieved both marks on **part 6aii** they then achieved at least one mark on this part of the question, with many achieving both marks and showing good knowledge and understanding of the reasons for controls.

(b) (i) This three-mark question asked candidates to draw a graph showing only the results for 'photographs with people' from Table 6.1. Candidates had to label the axes. This question caused very few issues for candidates with the vast majority getting the full three marks. Where candidates did not achieve full marks, it was due to the drawing of a histogram rather than a bar chart which meant that they could only achieve a maximum of two marks. Some candidates did not give units for the y axis but as long as their bars made sense in terms of the units used then the drawing of the bars could be credited. It is worth reminding candidates to ensure that they give units for both axes whenever drawing a graph.

(ii) This two-mark question asked candidates to describe the conclusion that can be made from the data in Table 6.1. To achieve two marks candidate responses needed to give a complete conclusion which talked about both types of personalities (R and S) explicitly. If a conclusion is given which is correct but more general i.e. that personality type affects the type of scene people prefer, then one mark could be given. Candidates performed well on this question with the vast majority showing they understood the difference between a result and a conclusion. Most candidates also incorporated both types of personality within their conclusion.

Where candidates performed less well it was due to them repeating results rather than giving a conclusion. For example, if candidate responses talk about personality R SCORING higher than personality S then this is a result and not a conclusion. However, if they did the above but then went on to say, 'this means that.' then they could still be fully credited.

Question 7

(a) (i) This two-mark question asked candidates to suggest one way that Mr Grainger can measure the success of the training. For candidate responses to achieve the two marks available they needed to suggest an appropriate way and then give further detail about their suggestion. The suggestions needed to be quantified in some way to achieve the two marks. Candidates found this question challenging, although the full range of marks was seen. Where candidates performed well, they usually suggested counting the number of rewards used, or how long it took the animal to enter the stable without rewards.

Where performance was limited it either was due to the measurement not being quantitative, or suggesting how many of the animals entered the stable. Unfortunately, in the scenario it does say that all animals do enter the stable during the training, so this was not creditworthy. It is worth reminding candidates to not just look at the question but to ensure they read the scenario carefully to ensure they understand the demands of the question prior to answering.

(ii) This two-mark question asked candidates to identify the type of data produced by the measure of success and justify their answer. For candidate responses to achieve the two marks available they needed to identify that it would produce quantitative data and then justify their decision i.e. that it is

© 2024

numerical data. Where candidates had successfully given a measure that was quantitative in **7ai**, almost all were able to access the two marks available on this question. However, if candidates had given a qualitative measure in **part 7aii** they could not access the full two marks, but they were awarded a mark if they successfully identified that their measure was qualitative.

(iii) This two-mark question asked candidates to explain one weakness the type of data produced by the measure of success given. For candidate responses to achieve the two marks available they needed to identify a weakness and then give more detail about the weakness identified. Candidates who had successfully given a measure that was quantitative in **7ai**, were often able to access at least one of the marks available on this question.

The most common response seen was the lack of detail given by quantitative data which was creditworthy. However, some of these candidate responses then struggled to give relevant detail to achieve the second mark. Often there would be statements such as 'will not know how the animals were thinking' which really is not relevant for non-human research. Many, however, did suggest that it does not give any detail about whether an animal was reluctant/showing distress which not only showed knowledge of the weaknesses of quantitative data but also some thoughtfulness about the study itself.

For those candidates who had given a qualitative measure, one mark was available for suggesting a weakness of that type of data, but two marks could not be achieved.

(b) (i) This two-mark question asked candidates to suggest two situational variables that could have affected the success of training. To achieve the two marks available candidates needed to suggest two unique situational variables which may have affected the results of the study in relation to the training. Candidate performance on this question was pleasing. Many candidates were able to suggest variables such as distraction of other people/animals, but by far the most common response was about the weather.

Where candidates did not achieve the two marks it was either due to repetition. Suggesting it was colder one day and hotter one day could only really achieve one mark as it was about temperature as a whole, or due to candidates suggesting participant variables such as the animals not liking the food given, or the animals not liking the food; both of which are participant variables and therefore could not achieve any marks.

(ii) This two-mark question asked candidates to explain why this situational variable may have caused differences in the success of training of the animals. To achieve the two marks candidate responses needed to choose one of the situational variables suggested in **part bi**, suggest why it would it make a difference and then give some further detail on that reason why.

Candidate performance on this question was directly linked to their success in answering **part bi**. If candidates had suggested participant variables in **part 7bi** it was unlikely they were going to achieve any marks in **part 7bii**. The most common response given for this question was related to the weather, with candidates suggesting that the time the animals took to get into the stable may have been altered by the weather i.e. if it was sunny they may not want to go inside, so they would have been given more rewards due to the weather not the training.

Question 8

(a) This two-mark question asked candidates to suggest two ways in which Chen's study is similar to the study by Piliavin et al. To achieve the two marks available candidates needed to suggest two separate ways that Chen's study was similar to Piliavin et al. Candidate performance on this question was pleasing, with the vast majority of candidates able to achieve at least one mark, with a significant number of those candidates achieving two. The most common responses here highlighted that both studies were about helping behaviour (which was the first mark) and that they both had victims (second mark).

Where candidates did not achieve both marks it was invariably due to giving responses that was generic and could apply to any study, such as they were both field experiments, or that they both included observation. This was the minority of candidates however, and this question showed the excellent knowledge candidates had about this particular study.

© 2024

(b) This two-mark question asked candidates to suggest two ways in which Chen's study is different from the study by Piliavin et al. To achieve two marks candidates needed to identify two explicit differences which included a comparison between the two studies. For example, candidates could achieve the mark if they said that one was on a subway but the other on a bus, but could not achieve the mark if they only said that Chen's study was on a bus. It is worth reminding candidates that if they are asked for a difference between two studies/theories then they should make a comparative point in order to be sure they will achieve the mark available.

Performance on this question was slightly below **Question 8a**, but in the main candidates were able to access at least one of the marks. The most common response was related to the type of transport (bus not a subway), or that Chen was looking at age alone not race and gender which Piliavin et al. did. Very few candidates were unable to access at least one of the marks, but where this did occur it was usually due to lack of knowledge of the study itself, or assumptions which were not present in the scenario, such as one was in New York and the other was not (the scenario does not say where the study was conducted).

(c) This two-mark question asked candidates to explain the sampling technique Chen is using. To achieve the two marks candidates had to firstly identify the correct sampling technique which was definitive (opportunity/convenience sampling) and then give further detail about why this was the case in THIS study. The second point needed to be linked to the study in some way to get the mark. Candidate performance on this question was, in the main, pleasing. Most candidates correctly identified that it was opportunity sampling, but less were able to achieve the second mark. The most common error was not providing a link to the study saying 'because they were there at the time' which is generic and therefore cannot be credited. Many candidates were able to link back to the study and therefore achieve the second mark.

Question 9

(a) This ten-mark essay question asked candidates to describe how Tanya could conduct a study using a questionnaire for teachers to investigate a variety of prosocial behaviours in children. Candidate responses showed the full range of marks, with some really nice responses at the higher mark range. Candidates did still find it a challenge to achieve the higher Levels (5) but there were a significant number of Level 3/4 responses which was pleasing to see.

Many candidates had a sound understanding of questionnaires and were able to make relevant decisions about the type of questions used, and how to interpret the data they gained. Where candidates struggled slightly was with the technique used to administer the questionnaire. Often the technique was implicit within the response rather than explained which often meant that candidates were unable to achieve the higher levels.

One of the elements that candidates needed to achieve to get into the higher levels was to give examples of questions and candidates excelled on this part of their response, with many giving relevant and accurate examples of the questions they could use, and the type of data that this could achieve which was pleasing to see.

Candidate responses within the lower mark range were often able to give a list of questions which they were going to use and were able to produce a basic procedure to follow. However, at this level, candidate responses would have significant gaps within the procedure which would mean that it would not be replicable. Candidate responses may also mistakenly suggest other research methods that they could use, such as observation which was not creditworthy. In addition, at this mark range some candidates mistakenly suggested that it would be the candidates themselves that were responding to the questionnaire and not the teachers. In this case, the response could still achieve marks for suggesting types of questions, and interpretation of data. but it was inevitable that these responses struggled to get higher than Level 1 or 2. At this mark range some candidates still talked about ethics and sampling which was not creditworthy.

Candidate responses within the higher mark ranges would be able to describe a procedure that would be replicable by other researchers. Most candidates at this level would suggest both open and closed questions, and then describe in detail examples of questions that could be used within the study. This would include answer choices for closed questions and the correct command words such as describe or explain for the open questions. Within this description candidates would highlight the type of data that would be produced by these types of questions. At this range, candidates would then go on to describe how their data could be interpreted such as averages and

© 2024

bar graphs for the quantitative data and looking for themes within qualitative data. What differentiated candidate responses at this level was whether they explicitly talked about the technique used to administer the questionnaire. Some candidates would describe handing the teacher the questionnaire at lunch time, and then gathering it in after completing it, some would suggest emailing the teacher so they could do it in their own time. These candidate responses would be examples of those who would achieve high Level 4, or even Level 5. Very few candidates at this mark range discussed anything about ethics or sampling which was pleasing to see.

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

(b) (i) This two-mark question asked candidates to describe one practical/methodological strength of the procedure you have described in your answer to **part (a)**. To achieve the two marks candidates needed to identify a strength of their procedure, and then provide further detail for that point. Candidate performance on this question was mixed. Where candidates achieved the two marks, they identified appropriate parts of their procedure such as the type of question (closed questions meaning quantitative data), standardisation of their procedure (same questions for everyone) and were able to give relevant detail for the second point (such as the fact that the same questions mean it is replicable to test reliability). The most common error in responses was a lack of detail meaning that often the second mark was not available. Unfortunately, when candidates had completed observations as part of their procedure, or they suggested the candidates themselves to answer the questionnaire they often would discuss these in this part of the question but inevitably this would not be creditworthy.

(ii) This two-mark question asked candidates to describe one practical/methodological weakness of the procedure you have described in your answer to **part (a)**. Similar to **part 9bi**, candidates needed to identify a potential weakness of their procedure, and then explain/exemplify why this would be a problem. Again, candidate performance on this question was mixed, and the clarity of responses was directly related to the detail they had given in their procedure. Sometimes candidates would discuss potential problems which were either incorrect (discussing issues of qualitative data when they had not used open questions) or related to observational methodology; neither of these would be creditworthy.

A number of candidates produced thoughtful responses for this question, with issues such as a lack of detail as they had only used closed questions, or emailing the questionnaires meaning that they could not be sure who had answered the questionnaires. Some candidates also discussed that the teacher may not have been at the school long and therefore may not know the children well so may just 'guess' their answers, which showed a good awareness of potential issues and was well worth the two marks available.

4Uadmission

© 2024

PSYCHOLOGY

Paper 9990/22
Research Methods 22

Key messages

- Candidates need to ensure that their responses are focused on the questions within the exam paper. There was more than one instance where it was clear that candidates had misread the question and provided responses which were not creditworthy.
- Candidates need to ensure that they understand the expectations for different command words used on the paper; for example, describe and explain. For describe, candidates need to ensure that they provide a sufficient number of unique points related to the marks allocated to the question whereas for explain, candidates need to identify a particular feature/concept/theory and then **link** their detail point to the feature/concept/theory they have identified. Often, explain questions had poorer outcomes in terms of marks as a candidate just described.
- Candidates need to ensure that they are able to define/outline key terms within the syllabus such as 'informed consent' and 'counterbalancing'.
- Candidates need to ensure that they link their answers to the information given in the stem if asked to do so. Often candidates showed excellent understanding of named issues/studies but then lost marks for giving generic responses.
- It is worth noting that candidate responses for the higher (6 and 10 mark) tariff questions showed good knowledge and understanding of correlations and observations. There were a number of thoughtful responses for the extended essay question, and candidates should be commended for their performance on these questions.

General comments

This was the first November series for the new Psychology syllabus. Paper 22 scripts provided the full range of marks, showing a high level of knowledge and understanding across many areas of the syllabus. Where performance was limited, it was due to a lack of knowledge of key terms, or a misunderstanding of the demands of the question. This was clear when looking at questions on counterbalancing **(4bi)**, repeated measures design (**4a** and **4bii**), and informed consent (**3a** and **3b**). Candidate responses also showed some gaps in knowledge when referring to Hassett et al. (monkey toy preferences), with responses showing some misunderstanding in relation to the methodology and findings of the study. Candidates did show their ability to structure six– and 10-mark responses (**5** and **9a**) with many able to produce thoughtful procedures which incorporated all of the elements required within the question. It is clear that centres have prepared candidates well for the exam. For future series, candidates need to ensure that they have a good understanding of command words, key research method terminology such as validity and reliability, and the studies which have been named on the syllabus.

Comments on specific questions

Section A

Question 1

This two-mark question required candidates to describe how obedience was measured in the study by Milgram. To achieve two marks candidates needed to show that they understood that it was the shock generator that was used to measure obedience, and that a specific participant's obedience was then measured by how high they went on that generator. A full range of marks was seen for this question, with many candidates showing good knowledge and understanding of the study. Where candidate responses did not achieve the two marks it was due to a misunderstanding of the question, with some suggesting that

© 2024

obedience was measured through observation/or the number of prods given by the experimenter; these were not creditworthy.

Question 2

(a) This two-mark question asked candidates to explain one similarity and one difference between a positive and negative correlation. For the similarity by far the most common response was the idea of a relationship between two variables; and indeed, the majority of candidates achieved this mark. The most common error on this part of the question was suggesting that there was an IV and a DV in a correlation. The difference was the stronger in terms of candidate responses, with the vast majority able to suggest a creditworthy difference. There were very few issues in terms of responses for this question, and candidate knowledge about correlations as a whole was pleasing.

(b) This two-mark question asked candidate to describe one example of a correlation from the Hassett et al. (monkey toy preferences) study. For two marks candidates must give an example of a type of correlation (positive and no correlation were creditworthy), and then give detail in relation to the Hassett et al. study. Many candidates were able to gain at least one mark for a type of correlation; with by far the most common answer being a positive correlation, but then struggled to give creditworthy information from the Hassett at al. study to exemplify their first point. Consequently, the most common mark on this question was one, with a minority of candidates getting the full two marks.

Candidates seem to struggle slightly with the findings of the Hassett et al. study and often gave responses which were either incorrect or did not actually show the correlation they identified. The results of this study are detailed and wide-ranging, therefore teachers should consider how the findings could be made clearer to candidates for future series.

Question 3

(a) This two-mark question provided some challenges for candidates. This question asked candidates to explain what is meant by informed consent. To achieve two marks candidates needed to suggest that participants needed to be given enough information about the study, and then asked if they would be willing to participate. Many responses did not achieve marks for this question due to candidates explaining 'consent' rather than 'informed consent'. Where candidates did not explain that participants would be given information about the studies aim, and only said about being willing to participate they could not achieve any marks as they did not answer the question. The other common error was tautological responses where candidates repeated the key words i.e. 'where participants were informed about the study and then asked to give their consent' would not be creditworthy.

(b) This two-mark question asked candidates to suggest how informed consent could have been obtained from the children in the study. To achieve two marks candidates needed to suggest a child friendly way of informing the children about the study, and then highlight that they would need to ask the child whether they wanted to be involved/play with the toys. Candidate responses on this question were mixed. As in **part (a)** many responses just concentrated on the idea of consent but struggled to suggest an appropriate way of telling children about the study therefore a minority of candidate responses were able to achieve the full 2 marks.

Question 4

(a) This two-mark question asked candidates to outline what is meant by a repeated measure design using an example from Daniel's study. To achieve two marks candidates needed to explain what is meant by repeated measures, and then show their understanding through an example. Most candidates were able to achieve at least one mark for this question, with a significant number of candidates achieving both marks available. Where performance was limited, it was usually due to candidates not making it clear that the same participants were in both levels of the IV/condition, with a minority wrongly suggesting that it was when participants do the experiment more than once. Most candidates were able to give an appropriate example from Daniel's study. The most common mark on this question was one.

(b) (i) This two-mark question asked candidates to outline what is meant by 'counterbalancing', using an example from Daniel's study. To achieve two marks candidates had to show their knowledge of counterbalancing and then exemplify this knowledge using the study provided. Candidates found

© 2024

this question challenging. Often it was clear that candidates knew the meaning of counterbalancing but struggled to put this is into a coherent outline on paper. Consequently, many candidate responses were muddled and unclear and therefore did not achieve any marks. Where candidate responses achieved marks, it was due to a clear outline of the changing of the order between sets of participants, or the use of the 'ABBA' technique which was often explained well. Where candidates achieved the mark for their understanding of counterbalancing, they often achieved the second mark for an appropriate example.

(ii) This two-mark question asked candidates to explain one strength of a repeated measures design. To achieve two marks candidates needed to identify a strength and then give further detail/elaboration of the strength identified. Performance on this question was mixed. The majority of candidates were able to identify a strength, with a reduction of individual differences or the use of less participants by far the most common responses. Where candidate responses lost marks, it was due to a lack of detail to exemplify this strength. Many responses just added 'as they were in both conditions', which was not enough for the second mark.

Question 5

This six-mark question required candidates to describe the techniques of paper and pencil, and online questionnaires. Awarded marks could come specific/accurate features of each technique and appropriate examples such as the IQ test from Baron-Cohen et al. being completed online. Candidate responses to this question were pleasing. At the lower end of the mark range, candidates usually achieved one or two marks for suggesting that online questionnaires were completed using the internet/email however they often struggled to give responses to 'paper and pencil' questionnaires which were not tautological i.e. using the words paper and pencil.

At the high end of the mark range candidates produced some thoughtful responses which looked at a variety of features of each type of questionnaire, and relevant studies which had used these methodologies successfully. It is worth noting that examples do not have to be from named studies, and there were some creative responses which used a variety of examples, which are creditworthy. Lastly, it is worth noting that responses cannot be credited twice for the same information no matter if it is used in different techniques, such as the use of both open and closed questions/obtaining qualitative and quantitative data.

Question 6

(a) This two- mark question asked candidates to identify the two independent variables (IVs) in Mohsin's study. The vast majority of candidates gained both of the marks for this question and no real issues were found. The most common error was candidates who only gave one IV across both parts of the question; writing healthy snacks in **part 1** and unhealthy snacks in **part 2**, meaning that they missed out on the second mark; however, this was a minority of responses.

(b) (i) This two-mark question asked candidates to state two other combinations that Mohsin should display in the additional weeks. The answer to this question was definitive and unfortunately most candidates found this question a challenge and did not achieve any marks. The most common error was that candidate responses would still combine unhealthy and healthy snacks but with blue and green packets. Although this of course was a different combination, it would not have been useful in this study and therefore achieved no marks. Some candidates gave some creative responses using a variety of different colours/food etc. but they needed to use the snack types/colours already stated so this again was not creditworthy.

(ii) This two-mark question asked candidates to explain why the two combinations stated in **part (b)(i)** would be useful. To achieve two marks candidates needed to refer back to the responses they gave in **part (i)** and explain why this was useful, and then elaborate/exemplify their point. As this question was explicitly linked to **part (b)(i)** candidates who got this part incorrect were unable to achieve any marks on this part. Where candidates did achieve marks on **part (i)**, the vast majority were able to get at least one mark for suggesting that it controls for food choice differences, but often struggled to give enough detail for the second mark.

(c) This one-mark question asked candidates to define the term 'population'. Responses needed to ensure that this was a definition relevant to psychology rather than a general definition such as 'all the people in one area'. The majority of candidates were able to achieve the mark available on this question by suggesting it was people who shared characteristics. There were a significant number

© 2024

of responses which gave a general definition however, therefore candidates need to be reminded to give definitions relevant to psychological processes rather than dictionary definitions.

(d) This two-mark question asked candidates to outline the population that Mohsin used in his study. To achieve two marks candidate responses needed to give two separate features of Mohsin's population such as supermarket shoppers/those buying food or snacks, those that live close to the supermarket/in his town. This question caused very few issues to candidates both at the bottom and top of the mark range. Most candidates were able to achieve at least one mark; usually for supermarket shoppers, with a significant majority achieving both marks.

Question 7

(a) (i) This one-mark question asked candidates to write one closed question that Jenny could use to investigate creativity. For candidate responses to achieve the mark available they needed to ensure that the question was related to the topic area being investigated, and ensure that they provided the answer choices for their question. This could be rating scales/yes or no/or providing alternative responses within the question. Most candidate responses did provide answer choices and therefore were able to achieve the mark. This was pleasing to see, and a vast improvement from previous series. Where candidate responses did not achieve the mark, it was usually due to a lack of answer choices, although a small minority of candidates still asked questions such as 'how old are you' (with answer choices), which was not creditworthy as it is not relevant to the topic area being researched.

(ii) This one-mark question asked candidates to write one open question that Jenny could use to investigate creativity. For candidate responses to achieve the mark available they needed to ensure that the question was truly an open question; using command words such as describe/explain/tell me why. It should be noted that how and what were not creditworthy as an open question. Candidates found this question slightly more challenging than the previous one; with many using what and how and not describe/explain etc. The majority of candidates were able to access the mark available.

(b) This three-mark question asked candidates to suggest why differences between the subjects they study could also explain the differences in doodling. To achieve the marks available for this question they either had to make one suggestion (such as suggesting that a subject may be boring) and then give two further points of detail (doodling to concentrate, if more interesting you'd pay attention more) OR give two suggestions and then a further point of detail for ONE of the suggestions. In the main, candidates performed well on this question, with the vast majority of candidates being able to access at least some of the marks, with many accessing all three. By far the most common suggestion was that a subject was boring, and how this will lead to lack of concentration and doodling. The most common error in responses was the lack of enough detail to achieve all three marks, but this was in the minority and there were some really thoughtful and insightful responses seen.

(c) This two-mark question asked candidates to suggest why doing an overt observation may not produce valid results. To achieve the two marks candidate responses needed to suggest one reason why the methodology chosen would not achieve valid results, and then give a further point of detail (which did not need to be related to the scenario). Candidates were often able to achieve at least one mark out of the two available; usually for the suggestion of demand characteristics, but then they did not give enough detail for the second mark to be awarded. Often candidate responses would suggest demand characteristics OR becoming aware of the aim, which was unfortunate as putting both together would have enabled them to achieve that second mark.

Question 8

(a) This two-mark question asked candidates to write axis headings for the *x*-axis and the *y*-axis on Fig 8.1. (a bar chart). This was a straightforward question for most candidates with many achieving the two marks available. Where candidates did not achieve the two marks it was usually due to mislabelling the x axis with some saying 'when they go to work' rather than 'day of the week', or just labelling the y axis as 'smiling at work' which is not quantified and therefore not creditworthy.

(b) This two-mark question asked candidates to explain one reason why it would have been better to have conducted the study for more than one week. To achieve the two marks candidate responses need to identify a reason for conducting the study for a longer period of time, and then exemplify

© 2024

this reason; this needed to be linked to the study in some way. Candidate responses to this question were, in the main, pleasing. The vast majority of candidates were able to get at least one mark for suggesting that this would improve generalisability (by far the most common response), validity etc. and many of these responses would then go on to exemplify this, such as saying there may have been something about that week that caused them to smile / not smile. The most common error in responses was the lack of detail given for the second mark or giving vague responses such as 'because it gives more data', which is not creditworthy.

(c) (i) This two-mark question asked candidates to suggest one other way that Parul could have measured happiness. To achieve two marks, candidates needed to identify an alternative way and then give some further detail. In the main, candidates often struggled to achieve the two marks on this question. The majority of candidates could identify an alternative way, such as laughing, positive body language or use of a questionnaire but many then were unable to give enough detail to achieve the second mark. Inevitably, it was the responses that suggested a questionnaire that achieved the two marks as they were able to suggest a scale that could be used which would then be their detail mark.

(ii) This two-mark question asked candidates to explain one advantage of the measure of happiness suggested in **part (c)(i)** compared with counting the number of colleagues smiling. To achieve the two marks candidates needed to explain an advantage, and then give some detail about the advantage identified. The success of this question inevitably depended on the success of their response in **part (c)(i)**. Those candidates who chose to suggest a questionnaire in **part (c)(i)** often were able to achieve the marks in **part (c)(ii)** through the suggestion of it being more objective/less researcher bias; however, those candidates who gave suggestions such as body language or length of talking between colleagues often struggled to suggest an advantage of this measurement and did not achieve marks in this part of the question. If candidates did not achieve marks in **part (c)(i)** it was unlikely they would achieve marks in **part (c)(ii)** although not impossible if the only issue on (**c)(i)** was the vagueness of the response rather than it being incorrect. With this type of question, it may be worth candidates looking at all parts of the question prior to answering the first part, as this may influence the type of answer they choose.

Question 9

(a) This ten-mark essay question asked candidates to describe how Chloë could conduct an observational study using a structured observation to investigate the social behaviour of adults in a town centre during their lunch break. Candidate responses showed the full range of marks, with some thoughtful responses at the higher mark range. Candidate responses found it a challenge to achieve the highest Level (5) but there were a significant number of Level 4 responses which was pleasing.

Many candidates had a sound understanding of observations and were able to make relevant decisions about whether it should be covert or overt, participant or non-participant. There was a lack of understanding about what was meant by a naturalistic observation and this often meant that candidate responses were not able to achieve the highest mark band. As the question already stated that the study should be structured, candidates needed to explain how data would be recorded in a structured observation i.e. using behavioural categories; and candidates excelled on this part of their response, with many giving relevant and accurate examples of the categories they could use, and some in the higher mark range discussing time and event sampling for their recording of data.

Candidate responses within the lower mark range were often able to give a list of behavioural categories which they were going to observe and were able to produce a basic procedure to follow. However, at this level, candidate responses would have significant gaps within the procedure which would mean that it would not be replicable. Candidate responses may also mistakenly suggest that they would use an unstructured observational methodology. This would not be creditworthy as the question required the research to be a structured observation. In addition, at this mark range some candidates suggested the use of questionnaires alongside observation, or explicitly used an experimental method with very little mention of observational methods which again is not correct.

Candidate responses within the higher mark ranges would be able to describe a procedure that would be replicable by other researchers. Most candidates at this level would suggest a covert observation, and then describe how this would be achieved. For example, they would suggest that the researcher should pose as a worker on lunch break and sit on a bench having lunch in an area

© 2024

that they could observe a wide range of people. The behavioural categories described at this mark range were thorough and, for the most part, fully operationalised with time and event sampling discussed and often justified. Most candidates at the higher mark range were able to show understanding of what is meant by a naturalistic observation although it was still slightly muddled at times. Candidates understanding of all other features of the observation were clear and explicit.

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

(b) (i) This two-mark question asked candidates to explain how one feature of the procedure they described in **part (a)** helps to make the study valid. To achieve the two marks candidates needed to identify a part of their procedure that helps validity, and then provide further explanation for that point. Candidate performance on this question was mixed. Where candidates achieved the two marks, they identified appropriate parts of their procedure such as the type of observation (covert meaning natural behaviour), operationalisation of their categories, and where the study was conducted (such as in a coffee shop/café) and were able to give relevant detail for the second point (such as the type of validity this would help/an explanation why this would help validity). The most common error in responses was a lack of detail meaning that often the second mark was not available. Unfortunately, when candidates had completed questionnaires/experiments as part of their procedure they often would discuss these in this part of the question but as this was the wrong methodology this would not be creditworthy.

(ii) This two-mark question asked candidates to explain how one feature of the procedure they described in **part (a)** could be a problem for the validity of the study. Similar to **part 9(b)(i)**, candidates needed to identify a potential problem within their procedure, and then explain/exemplify why this would be a problem. Again, candidate performance on this question was mixed, and the clarity of responses was directly related to the detail they have given in their procedure. Sometimes candidates would discuss potential problems which were either incorrect (discussing issues of qualitative data when they had not mentioned this within their procedure) or related to non-observational methodology; neither of these would be creditworthy. Several candidates produced thoughtful responses for this question, with issues such as the observer being spotted 'observing' changing behaviour, a lack of operationalised categories, and situational factors such as the weather, time of year etc.

4Uadmission

© 2024

PSYCHOLOGY

Paper 9990/23
Research Methods 23

Key messages

- Candidates need to ensure that their responses are focused on the questions within the exam paper. There was more than one instance where it was clear that candidates had misread the question and provided responses which were not creditworthy.
- Candidates need to ensure that they understand the expectations for different command words used on the paper; for example, describe and explain. For describe, candidates need to ensure that they provide a sufficient number of unique points related to the marks allocated to the question, whereas for explain candidates need to identify a particular feature/concept/theory and then **link** their detail point to the feature/concept/theory they have identified. Often, explain questions had poorer outcomes in terms of marks as a candidate just described.
- Candidates need to ensure that they can define/outline key terms within the syllabus such as 'random sampling' and know the difference between terms such as qualitative and quantitative data. This is also important in questions where key terminology is used. It was clear that in some questions which talked about terms such as operationalisation, there was a lack of understanding about what this word actually meant.
- Candidates need to ensure that they link their answers to the information given in the stem if asked to do so. Often candidates showed excellent understanding of named issues/studies but then lost marks for giving generic responses.
- It is worth noting that candidate responses for the 10-mark extended response question showed good knowledge and understanding of interviews. There were several thoughtful responses for this extended response question, and candidates should be commended for their performance.

General comments

This was the first November series for the new Psychology syllabus. Paper 23 scripts provided the full range of marks, showing a good level of knowledge and understanding across many areas of the syllabus. Where performance was limited, it was due to a lack of knowledge of key terms, or a misunderstanding of the demands of the question. This was clear when looking at questions on random sampling (**1**), operationalisation **(8b)**, and the use of stooges (**4**). Candidate responses showed good knowledge and understanding of key studies such as when referring to the Fagen et al. (elephant learning) study, and the Perry et al. (personal space) study, although responses did show some misunderstanding in relation to the Pozzulo et al. (line ups) study and how the mean was calculated. Candidates did show their ability to structure 10-mark responses **(10a)** with many able to produce thoughtful procedures which incorporated all the elements required within the question. It is clear that centres have prepared candidates well for the exam. For future series, candidates need to ensure that they have a good understanding of command words, key research method terminology such as validity and reliability, and the studies which have been named on the syllabus.

Comments on specific questions

Section A

Question 1

This one-mark question required candidates to outline what is meant by the term 'random sampling'. To achieve the two marks candidates needed to give an accurate explanation of the term, such as 'participants having an equal chance of being selected'. If a candidate outlined how a random sample is obtained, then credit could also be given.

© 2024

Candidate performance on this question was pleasing, with the majority of candidates were able to achieve the mark in one of the two ways suggested. When candidates did not achieve the mark it was due to tautological responses, such as 'when the participants are chosen at random', or a confusion with opportunity sampling, however this was a minority of candidates.

Question 2

This two-mark question asked candidates to identify the dependent variable (DV) in a study, and to justify their answer. To achieve the two marks candidate responses needed to identify the correct dependent variable (happiness) and then give a justification of why this is the case. The vast majority of candidate responses gave the correct dependent variable, with most of these candidates able to give an appropriate justification, with the most common response being that it was the variable being measured. Where candidates did not achieve the two marks it was invariably due to confusion between the independent and dependent variable which meant that these candidates did not achieve any marks.

Question 3

(a) This one-mark question asked candidates to define what is meant by 'qualitative data'. There were no major issues with this question with almost all candidates able to show accurate understanding of the term. Where candidates did not achieve the one mark available it was due to them suggesting that qualitative data is 'words' which is not creditworthy.

(b) This two-mark question asked candidates to define what is meant by 'quantitative data' and include an example from Fagen et al.

Candidate performance on this question was pleasing. Most candidates were able to get at least one mark from this question with an appropriate definition of quantitative data, with the most common responses being the idea that it was numerical. It was also pleasing to see that most of these responses were able to then go on and give an appropriate example from Fagen et al. such as time taken/percentage success of tasks.

Where candidates did not get the second mark it was due to the example given not being quantitative, or a lack of understanding of the Fagen et al. study. However, this was a minority of candidates.

Question 4

(a) This two-mark question asked candidates to suggest one way to improve the ecological validity of this study in relation to the victims in the Piliavin et al. study. To achieve the two marks candidate responses needed to suggest a way which would improve ecological validity and then provide a link back to the Piliavin et al. study.

Candidate performance on this question was mixed. Most candidates were able to show clear knowledge and understanding of the study and the term, but some would then go on to give unviable, or unethical suggestions which could not gain credit. The most common response which achieved the two marks was the idea of using both male and female 'victims', as females have different characteristics which may affect helping behaviour, or making the 'victim' more realistic, such as actually putting their leg in plaster.

Where candidate performance was limited it was due to responses such as, 'using a person who was actually drunk' or 'giving a person real alcohol so they were actually drunk'. Unfortunately, this is not an ethical way of conducting a study of this nature so these sorts of responses cannot be credited. It is worth noting that in relation to the 'ill' person it would be perfectly ethical to use a person whose leg was already in plaster, or who had some mobility issues in real life, as this would be ethical as you are not causing their broken leg!

(b) (i) This two-mark question asked candidates to suggest one reason why measuring responses to imagined people is more ethical than measuring responses to stooges. To achieve two marks candidates needed to give one reason why it is more ethical and then ensure that this response was linked to the study in question. Candidate performance on this question was directly related to their depth of understanding of the Perry et al. study.

© 2024

Where candidates had good understanding of the study they were invariably able to achieve at least one, and in many cases two, marks. The most common response was that it was less distressing (ethical point) as you do not actually have a stranger come close to you (link to the study).

Where candidates did not achieve the two marks available it was due to a lack of knowledge of the study itself, or very vague responses such as 'it was less ethical due to the stranger coming close to you' – although this response could get a basic one mark, they have not given a specific ethical issue, such less chance of possible harm/distress/less deception etc., therefore cannot achieve the full two marks.

(ii) This two-mark questions asked candidates to suggest one reason why measuring responses to stooges is more practical than measuring responses to imagined people. To achieve two marks candidates needed to give one PRACTICAL reason and then link it to the study in question.

Question 5

This two-mark question asked candidates to explain how the researchers calculated the mean in the Pozzulo et al. study. To achieve two marks candidates needed to give a full and correct explanation of how the mean was calculated, whereas a partial explanation (such as describing how a generic mean is calculated) would achieve one mark.

Candidates found this question challenging with the vast majority of candidates achieving only the one mark. It was clear that candidates had good knowledge of how to calculate a mean score, but many were not able to place this in the context of the Pozzulo et al. study.

The correct response for this question was the number of correct identifications from the children (which most actually got) divided by the number of possible identifications (the number of targets shown) – it was this last part that many candidates did not get, with many suggesting that it was the total number of children. This was not correct, but as it shows clear understanding of how to calculate a 'generic' mean then one mark was awarded.

Candidates should be aware that questions about the calculations of results, such as the above, are legitimate questions for all studies named on the specification so care should be taken to ensure that the methodology used is understood prior to the examination.

Question 6

This six-mark question required candidates to describe laboratory experiments and field experiments, using any example(s). Awarded marks could come from accurate descriptions of each term, further detail about the term and appropriate examples such as showing knowledge of the IVs and DVs of studies on the syllabus, such as Piliavin et al. study having cane and drunk victim as the IV and helping behaviour as the DV. Candidates performed exceptionally well on this question, with most having excellent knowledge and understanding of both types of experiments and able to give appropriate examples. The majority of candidates achieved most, if not all, the marks on this question.

At the lower end of the mark range, candidates usually achieved one or two marks for suggesting that a laboratory experiment is highly controlled with an independent and a dependent variable. They may be able to give an appropriate example such as from the Andrade study. Candidate explanations of field experiments were weaker and often were limited to suggesting that it was based in 'real life' situations, linking this to the Piliavin et al. study. Even at this lower end of the mark range, candidates were achieving 3 or 4 marks.

At the high end of the mark range candidates produced some thoughtful responses and the vast majority were able to give detailed explanations of both field and laboratory experiments, highlighting a number of features. The examples given were appropriate and explicitly linked to the question. These candidate responses often achieved full marks on this question.

Question 7

(a) This four-mark question asked candidates to suggest two techniques, other than interviews, that Daiyu could use in her case study to collect data about the child and justify their answers. Candidate performance on this question was mixed.

© 2024

Most candidates were able to name two alternative techniques, such as observations and questionnaires, however very few were able to justify their response therefore two was the most common mark. The most common error with this question was that candidates would justify their response using a positive aspect of the methodology (such as questionnaires can be more reliable) rather than justifying the method through an aspect of the child's autism (that autistic children may struggle with communicating verbally). Candidates to need to ensure that they understand the demands of the question prior to answering it.

(b) This four-mark question asked candidates to suggest two pieces of information about the child's interactions that would be useful for Daiyu to collect in her case study and justify their answers. Candidates found this question challenging with many misunderstanding the demands of the question.

For those candidates who did achieve four marks, the most common response was the idea of looking at differences between how the child reacts to children and adults and not making eye contact with others. Candidates then justified their responses thoughtfully, showing insight about how autistic children struggle with eye contact, and how autistic children may communicate better with children than adults.

For candidates who did not achieve the marks it was usually due to very vague responses which may have talked about differences in communication but did not specify who with and therefore did not give enough information to gain credit. At this mark range there were very few relevant justifications given.

(c) This one-mark question asked candidates to outline one practical reason why obtaining information from the child could be difficult. This question did not produce many issues with most candidates able to achieve the mark. The most common response was the increased difficulty communicating with children in general, and that some children with autism struggle more to communicate effectively. When candidates did not achieve the mark it was due to them giving ethical rather than practical reasons, such as not being able to gain consent from children.

Question 8

(a) This two-mark question asked candidates to explain the type of observation that Hazel plans to use because she is only recording whistling, singing and shouting. Candidate performance on this question was pleasing with most candidates able to achieve at least one mark for identifying that this was a structured observation.

Where candidate performance was more limited was when providing an explanation for their choice. Although a majority of candidates were still able to talk about categories/fixed behaviours, some suggested it was due to limited time or performing the observation in a laboratory which is not creditworthy.

(b) This one-mark question asked candidates to outline how Hazel could operationalise **one** of the three behaviours she plans to record. This question was surprisingly challenging for candidates, with many looking at timings, or providing repetitive responses such as 'the type of songs sung'.

Candidates who did achieve the mark available usually operationalised shouting as 'talking very loudly' which was creditworthy, but again some suggested 'how loud they shout' which was again repetitive. Candidates need to ensure that they understand key terminology such as operationalisation and how that fits into the demands of the specific question.

(c) (i) This two-mark question asked candidates to suggest one reason why it would be more ethical to be an overt observer than a covert observer in this study. To achieve the two marks candidates needed to suggest an appropriate ethical reason and link this ethical reason back to the study.

Most candidates were able to achieve at least one mark for the ethical issue, with the most common response being right to withdraw, but often did not get the second mark as their response was generic. For example, they would say 'because they may not want to be observed in public'; this was not specific enough for the link with something such as 'because they can tell the observer they do not want their **children** to be watched.

© 2024

Candidates need to ensure that when a question says 'this study' in the question that they link their response **explicitly** to the scenario given.

(ii) This two-mark question asked candidates to suggest one reason why it would be more practical to be a covert observer than an overt observer in this study. Like in **part 8(c)(i)**, to achieve the two marks candidates needed to suggest one practical reason why covert observation would be more practical and then link their response to the study. Candidate performance on this question was slightly better than **part 8(c)(i)**, with most candidates able to suggest a practical reason, and many linking to the question well.

The most common response was the lack of demand characteristics / social desirability and then a link to changing the amount they whistled if they knew they were being observed. Where candidates did not achieve the two marks it was usually due to a generic response, such as saying 'because they may change their behaviour' which is not linked enough for the second mark.

Question 9

(a) (i) This two-mark question asked candidates to suggest one way that Hudson could measure sleepiness, other than an interview. For candidates to achieve the two marks available they need to suggest a way to measure sleepiness and then provide some detail about that suggestion. Candidate performance on this question was pleasing, with most candidates able to suggest either an observation or questionnaire and therefore to achieve at least one mark.

Candidates who suggested a questionnaire were more likely to achieve the second mark, suggestion a question asking how sleepy a person feels on a Likert scale. Where candidates struggled slightly is when suggesting the use of an observation as they did not give a measurable category giving vague suggestions such as 'how sleepy they look'.

Candidates also used digital technology as a way of measuring sleepiness, and they showed their excellent knowledge of health apps / smart watches and often gained the second mark.

(ii) This two-mark question asked candidates to explain one strength of the measure of sleepiness they suggested in **part (a)(i)**. To achieve the two marks candidates needed to suggest one strength and then provide some detail of this strength.

Candidate performance on this question was directly related to how they answered **part 9(a)(ii)**. If candidates suggested digital technology, they were often able to give an appropriate strength such as objectivity of the data given rather than an observation where the researcher has to interpret behaviour. As seen in this response, it was often easier for candidates to achieve the second mark if they make a comparison to another method, and this is something worth highlighting to candidates.

Where candidates did not achieve the two marks, it was due to not giving enough detail from their strength to get the second mark, often just giving a generic strength of the method suggested and little more.

(iii) This two-mark question asked candidates to explain one weakness of the measure of sleepiness they suggested in **part (a)(i)**. To achieve the two marks candidates needed to suggest one weakness and then provide some detail of this weakness.

Candidate performance on this question was directly related to how they answered **part 9(a)(ii)**. If candidates suggested digital technology or questionnaires, they were often able to give an appropriate weakness such as the data being quantitative therefore there is no detail of why they are sleepy or people not wanting to admit they are sleepy leading to socially desirable responses. As seen in the above response, it was often easier for candidates to achieve the second mark in this question than **part 9(a)(ii)** as often it could be gained without a comparison, therefore performance overall was slightly stronger than in **part 9(a)(ii)**.

Where candidates did not achieve the two marks, it was due to not giving enough detail from their weakness to get the second mark, often just giving a generic weakness of the method suggested and little more.

© 2024

(b) This four-mark question asked candidates to sketch a scatter graph using the axes provided, to show what the pattern of results would look like if Hudson found a negative correlation. Candidates needed to label the axes.

Candidates found this question fairly straightforward with many getting the full four marks. Where candidates did not achieve the full four marks it was invariably due to candidates not putting the unit of measurement (hours) in the time since the last meal. Even so, candidates would normally get 3 out of the 4 marks available for this question.

Question 10

(a) This ten-mark essay question asked candidates to describe how Zeta could conduct a study using an interview to investigate the different ways that people respond to advertisements for different products. Candidate responses showed the full range of marks, with some really nice responses at the higher mark range. Candidates did still find it a challenge to achieve the higher Level (5) but there were a significant number of Level 3/4 responses which was pleasing to see.

Many candidates had a sound understanding of interviews and were able to make relevant decisions about the type of questions used, and how to interpret the data they gained. Where candidates struggled slightly was with the format of the interview which was slightly surprising. Often the format was implicit within the response rather than explicit which often meant that candidates were unable to achieve the highest levels. One of the elements that candidates needed to achieve to get into the higher levels was to give examples of questions and candidates excelled on this part of their response, with many giving relevant and accurate examples of the questions they could use, and the type of data that this could achieve which was pleasing to see.

Candidate responses within the lower mark range were often able to give a list of questions which they were going to use and were able to produce a basic procedure to follow. However, at this level, candidate responses would have significant gaps within the procedure which would mean that it would not be replicable. Candidate responses may also mistakenly suggest other research methods that they could use. For example, some candidates mistakenly suggested independent and dependent variables and detailed experimental procedures. In this case, the response could still achieve marks for suggesting types of questions and interpretation of data, but it was inevitable that these responses struggled to get higher than Level 1 or 2. At this mark range some candidates still talked about ethics and sampling which was not creditworthy.

Candidate responses within the higher mark ranges would be able to describe a procedure that would be replicable by other researchers. Most candidates at this level would suggest both open and closed questions, and then describe in detail examples of questions that could be used within the study. This would include answer choices for closed questions and the correct command words such as describe or explain for the open questions. Within this description candidates would highlight the type of data that would be produced by these types of questions. At this range, candidates would then go on to describe how their data could be interpreted such as averages and bar graphs for the quantitative data and looking for themes within qualitative data. What differentiated candidate responses at this level was whether they explicitly talked about the technique used for the interview, such as structured/unstructured and linked this in with the types of question structure they were using. These candidate responses would be examples of those who would achieve high Level 4, or even Level 5. Very few candidates at this mark range discussed anything about ethics or sampling which was pleasing to see.

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

(b) (i) This two-mark question asked candidates to explain how one part of the procedure they described in **part (a)** helps to make the study reliable. To achieve the two marks candidates needed to identify a part of their procedure that helps reliability and then provide further explanation for that point.

Candidate performance on this question was mixed. Where candidates achieved the two marks, they identified appropriate parts of their procedure such as using the same questions with everyone and the same procedure and were able to give relevant detail for the second point (such as the fact that the same questions mean it is replicable to test reliability).

© 2024

The most common error in responses was a lack of detail meaning that often the second mark was not available. Unfortunately, when candidates had completed experiments as part of their procedure, they would often discuss these in this part of the question but inevitably this would not be creditworthy.

(ii) This two-mark question asked candidates to explain how one part of the procedure they described in **part (a)** could be a problem for the reliability of the study. Like **part 10(b)(i)**, candidates needed to identify a part of their procedure that may cause a problem for reliability and then explain/exemplify with this point. Again, candidate performance on this question was mixed, and the clarity of responses was directly related to the detail they gave in their procedure

A number of candidates produced thoughtful responses for this question, with issues such a using unstructured interviews so would use different questions for each participant (meaning that there was no consistency in the approach to the interview), and then would go on to say how this made the interview difficult to replicate to test of reliability, or that there may be extraneous variables affecting reliability if they did not keep the procedure standardised.

The most common error in responses was a lack of detail meaning that often the second mark was not available. As in **part 10(b)(i)**, where candidates had completed experiments as part of their procedure, they would often discuss these in this part of the question but inevitably this would not be creditworthy.

4Uadmission
SINCE 2020

© 2024

PSYCHOLOGY

Paper 9990/31

Specialist Options: Approaches, Issues and Debates 31

Key messages

Questions 1, 3(a), 5, 7(a), 9, 11(a), 13, 15(ai), 15(aii) –

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

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

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

Questions 2, 6, 10 and 14 –

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

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

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

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

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

This question will always ask the candidate to evaluate the studies, theories, characteristics/explanations/treatments of disorders or techniques described in **part (a)** of the question. The

© 2024

response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. However, most responses that evaluated using two issues in this exam, achieved in the lower bands due to the response being superficial and often with little analysis. Some responses that considered three issues tended to achieve higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the studies, theories, characteristics/explanations/treatments of disorders or techniques described in the **part (a)** of the answer.

The candidate must also provide some form of analysis to access Level 3 and above. This could be done by discussing the strengths and weaknesses of the issue being considered, presenting a counter-argument to the issue under discussion or comparing the issue between two studies and/or theories. The response needs to explain the comparison/strength/weakness or counter-argument with examples from **part (a)** of the question. It was common for responses to state that two theories, for example, were 'similar' or 'in contrast' for an issue without any explanation as to why they could be compared in this way. This is limited analysis. A conclusion at the end of each issue would be helpful in order to show excellent understanding of the issue under discussion. In order to achieve the requirements of the Level 4 and 5 descriptors, it would be best to structure the response by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure the answer covers this requirement of the question.

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

General comments

The marks achieved by candidates for this section of the 9990 specification achieved across the full range of the mark band which was very pleasing to see. Some candidates were well prepared for the exam and showed good knowledge, understanding, application and evaluation throughout their responses. A significant number of candidates were not as well prepared and showed limited knowledge and understanding with brief, superficial and sometimes anecdotal responses. These candidates often had limited evaluation and application skills.

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

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

Comments on specific questions

Clinical Psychology

Question 1

The responses to this question covered the full range of marks. Responses that achieved 3-4 marks often referred to the process of systematic desensitisation by outlining the steps with accurate terminology. Common suggestions for contextualised responses involved Jude looking first at a picture of food followed by exposing Jude to real food eventually. Many candidates were able to outline the fear hierarchy and muscle relaxation. Weaker responses confused systematic desensitisation with other techniques. It was common for responses to fail to explain how the gradual exposure would occur so they achieved fewer marks overall.

Question 2

(a) There were some good responses that were able to clearly outline the reductionism side of the reductionism versus holism debate. For example, stating that it is explaining psychological phenomena by breaking it down into smaller component parts. One-mark responses often outlined that reductionism is to do with an explanation or concept that focuses on one thing/ignores other

© 2024

factors. Some responses were too specific, using definitions with psychological disorders/diseases as part of the definition which was not creditworthy.

(b) The majority of responses were able to explain the biochemical explanation for schizophrenia through outlining the dopamine hypothesis. However, only a handful were able to explain why this was reductionist (even when they had successfully defined reductionism in **part (a)).** Most attempted to explain why it was reductionist but often just stated it ignored other explanations, or it ignored cognitive explanations which was not creditworthy. Those that did achieve full marks were able to explain in some detail exactly what was ignored by this explanation such as lack of self-monitoring leading to hallucinations.

Question 3

(a) Responses to this question covered the full range of the mark scheme. Those that achieved 3-4 marks often were able to explain 'why' the Doctor would want to use the GAD-7. Common responses included that it is objective, quick and easy, to determine the severity of her anxiety or to see what appropriate treatment can be given. The best responses were able to support their statements with the GAD-7 e.g. quick and easy because it has only 7 items, or if to find out the severity of her anxiety due to her showing one of the symptoms (trouble relaxing) in the GAD-7. Weaker responses lacked any explanation as to why the doctor would want to use the tool. Some responses mentioned just 'quick and easy' without saying why it's quick and easy and achieved 1 mark for this reason given. A number of responses confused the measure (GAD-7) with Generalised Anxiety Disorder itself which was not creditworthy.

(b) There were some full mark responses to this question that could relate the weakness back to Aisha or the anxiety disorder. Those that did not achieve full marks lacked this link. Common weaknesses were that it lacks any in-depth explanation as to finding out the 'why' Aisha has anxiety. Other common responses were that the GAD-7 is a self-report measure which can lead to social desirability or Aisha lying. However, some responses gave incorrect weaknesses which received no credit including how the questions might be confusing and Aisha may not understand the questions. It should be noted the questions on the GAD-7 are very simple to understand. Another weakness that was not creditworthy included that the GAD-7 cannot be generalized to other phobias/disorders.

Question 4

(a) Responses varied for this question and covered the full range of the marks available. Level 3 responses were able to outline the study in detail, covering the participants, procedure (including one or more of the self-reports used to collect data) and results of the study as outlined in the mark scheme. Some responses were able to quote percentages of the results from the study although this was not a requirement for full marks. Weaker responses included fewer details of the study. A minority of responses confused this key study with another, but there were a significant number of responses that simply stated that it was a study comparing CBT face-to-face with that by telephone but with no details at all. Neither of these types of responses were creditworthy.

(b) The marks for this part of the question did cover the full range of marks available with the most frequent levels awarded being Level 2 and 3. Those that achieved Level 3 and above structured their response issue by issue, and often started with the named issue of reliability, along with evidence from the study in **part (a)** and analysis. Apart from the named issues, other popular issues covered were individual versus situational, quantitative data, validity, generalisability and application to everyday life. Popular examples for the named issue included how the timings of the telephone calls were standardised, and that all participants were given the same questionnaire and the sessions in each condition were conducted by the same therapist. Good responses were also able to outline how the study may not have reliability due to the lack of control on the surroundings of the telephone group. Another common response was that each individual's responses were different thus the way that CBT was conducted would not have been the same for every participant.

Common applications to everyday life were that the use of a telephone is less time consuming and can save cost on travel for therapy. Generalisability was also a popular point as participants were from the UK so it's difficult to generalise to other populations. Other responses like individual versus situational mentioned that the study is situational due to the outcome of the therapy sessions being dependent on whether they are in the telephone group or face to face group, yet also individual as each participant gives a unique response during CBT. Other responses covered

© 2024

the strength of using quantitative data from the questionnaires as it allowed for objectivity and comparisons.

Weaker responses achieving Level 1 or Level 2 did not contextualise their response. While some of the evaluation points were valid, the lack of context prevented responses from achieving a higher band. Some who wrote about a debate such as determinism versus free-will did not explain how the study supported the relevant side of the debate. Some provided too many issues with no depth in explaining why.

Consumer Psychology

Question 5

There were several good responses to this question and some achieved full marks. Full mark responses were able to outline two of the effects of changing to freeform layout (engages customer so longer browsing time, more enjoyable to use) and put this into the context of an online grocery retailer. Some outlined how the change might affect the customer such as being confused at the new layout which was also creditworthy. Weaker responses sometimes outlined one effect in context or were too brief (simply stating, for example, that it would be 'more entertaining'). A number of responses described customer behaviour in a physical rather than virtual store which was not creditworthy to achieve full marks.

Question 6

(a) Good responses were able to outline an application of using children e.g. having them watch an advertisement of a product aimed at children versus adults and seeing which brands they recognise. Some outlined showing children a few brands through a game before testing them to see which brand they are likely to remember or outlining the method from the Fischer et al. study. Some were able to relate it back to brand awareness e.g. by doing these it could help researchers gain insight to children's preference which will shape their marketing strategy. Most made some attempt at this question but lacked clarity as to why it would be suitable for children and this was frequently not creditworthy. A significant minority gave a description of the Auty and Lewis study, which is about product placement and not brand awareness so also not creditworthy.

(b) Some responses were able to gain full marks. These types of responses were able to outline a strength e.g. children are in critical stages of cognitive development which reveasl how branding shapes them over time/children are future consumers/children are less likely to have demand characteristics. The strongest responses were able to relate it back to brand awareness e.g. guide marketing for young audiences who will influence their parents' purchases. Weaker responses often lacked clarity and simply stated, for example, it helps understand the impact of advertising on children, without saying how. A response that was not creditworthy was stating that children can be manipulated very easily so companies could exploit this (suggesting unethical methods).

Question 7

(a) There were a number of full mark responses with two clear suggestions of how background music could improve the atmosphere. The most common responses referred to genre of music (almost always classical) and volume. Those that suggested classical music stated it would increase spending and referred to the North et al. study's results. Responses that could not achieve full marks often did not link to how the feature of music would affect the customer/atmosphere.

(b) Stronger, full mark response described how the genre of music needed to match the atmosphere of the restaurant and the effect this would have if this did not happen. Other common responses included that the music may distract customers and make them uncomfortable. Some responses talked about how the volume of music may be too loud and cause customers to leave or that the customers may not like the music/genre of music. There were a few novel suggestions including the cost of installing the music system in the restaurant. Responses that could not achieve full credit often did not link the problem back to the customer. Some described how the loudness of the music may affect the taste of the food, which refers to *noise* (rather than music in research) which was not creditworthy.

© 2024

Question 8

(a) There were many good, Level 3 responses to this question which included an outline Hall's four zones with the inclusion of a study. The vast majority summarised the Robson et al. study on table spacing in a restaurant. Candidates who knew about the overload, arousal and behaviour constraints could provide examples and relate that to restaurant tables. Weaker responses often gave an outline of Hall's four zones but the description of the Robson et al. study lacked detail and many very briefly outlined overload, arousal and behaviour constraint. Some responses did not include overload, arousal and behavioural constraint which limited the mark awarded to Level 2 at best. There were some very weak descriptions for this question with responses showing very little knowledge of the topic area and achieving either Level 1 or 0 marks for this reason.

(b) There were some Level 3 and above responses to this question. Most responses evaluated using the named issue of cultural differences with good examples given about how these differences might affect personal space at tables. Other common issues were application to everyday life, individual versus situational, generalisability (with reference to the Robson et al. study) and reductionism versus holism.

Weaker responses lacked depth in their discussion of cultural differences and simply stated that it could not be generalised outside of Western culture with no examples given which limited the marks awarded to Level 1 or Level 2. There was some confusion in weaker responses about the Robson et al. study with many incorrectly stating that it took place in a restaurant (when it took place as an online survey) so no marks were awarded for the discussion around good ecological validity. Weaker responses also provided a long list of evaluation issues with just a vague connection to the study/theory outlined in **part (a)**.

Health Psychology

Question 9

There were many strong responses to this question achieving 3-4 marks. Good responses made some good, clear suggestions (usually more than one). Strong responses suggested the use of mild fear arousal and described how pictures could be shown of the result of diabetes, obesity and/or heart disease. Others outlined how the teacher could use role models (similar to the Tapper et al. study) and rewards to encourage healthy eating. It was less common for responses to suggest how there could be a reduction in the eating of unhealthy foods and therefore full marks could not be awarded. Weaker responses did not make it clear that their suggestions would be applicable to the teacher's class (the responses could be for a whole school). Other weaker suggestions gave a very brief outline of their idea. Some suggestions were not practical for a teacher in a school or even possible (such as increasing the cost of unhealthy food or changing how it is advertised) and these were not creditworthy.

Question 10

(a) The majority of responses were able to outline what is meant by freewill. An example of a strong response is 'freewill is the individual's own autonomy and ability to choose to behave as they desire. For example, positive psychology states that individuals who have a happy life will choose to focus on their virtues. As they can choose this focus on their virtues, they can choose to be happy or decide against it as per their own accord.' Weaker responses did not provide an example from positive psychology. Some examples simply just said to 'think positively' as an example which was not creditworthy. Some responses gave a tautological definition which could not achieve credit.

(b) There were a few good responses to this question. Responses that achieved credit for this question often used examples from the Shoshani and Steinmetz study mentioning how candidates from single-parent / low-income families find it more difficult to achieve one of the lives due to their situation being determined by socioeconomic factors. Or the response used an example from meaningful life such as being able to do charitable work required the person to have access to this in their personal life which was not in their control. Many responses were not creditworthy due to lack of any reference to positive psychology and just stated, for example, that some people would never have good mental health (which clearly goes against the premise of positive psychology). Many responses gave a reductionist explanation instead of a deterministic one which also was not creditworthy.

© 2024

Question 11

(a) There were many good responses to this question. Most were able to outline at least one biological measure. Common responses were salivary cortisol, measuring heart rate using a pulse oximeter or measuring blood pressure. Many could achieve full marks for one or more of the biological measures by explaining how it would be used to make comparisons between home and work and/or what a high or low result on the measure meant in terms of stress. Lower mark responses just identified the measure. A significant number of responses suggested the use of an fMRI. This would clearly be unsuitable because, although such a device could be used as a biological measure of stress, fMRI machines are not portable devices (and could not be present in Nadim's work and home). A minority suggested non-biological measures of stress such as semi-structured interviews which was not creditworthy.

(b) There were some full mark responses to this question that explained the weakness of one of the measures from **part (a)** and put this into the context of measuring stress. The most common responses were that the measure may not truly reflect Nadim's stress levels and that his stress levels could be caused due to other named examples. Some candidates who talked about taking blood cortisol levels mentioned it will be troublesome for Nadim to take a blood sample at his office. No credit was given if the response explained that the test might not be administered correctly. If candidates evaluated the use of an fMRI (due to its unsuitability in **part (a)**), this was not creditworthy either.

Question 12

(a) The responses to this question covered the full range of the mark scheme. Better responses gave clear and often detailed and accurate descriptions of the Health Belief Model and reasons for non-adherence, with the strongest frequently giving details of the Laba et al. study. Weaker responses were able to achieve marks from writing about their knowledge of the Health Belief Model. Fewer responses were able to give a detailed outline of the rational non-adherence. Some only identified factors of the Health Belief Model and did not elaborate further on how these were barriers. Some responses were brief with short outlines of the Laba et al study. There were a significant number of candidates who did not know any content about rational non-adherence. There were also several responses which outlined the effect of doctor's clothing and tone of voice which was not creditworthy.

(b) The marks for this question covered the full range of marks. Most responses attempted the named issue of applications to everyday life. Common applications included using knowledge of the health belief model to lower medicine cost, do promotional campaigns on the benefits of medicine, and educating the public on the severity of certain illnesses. Another common issue was the individual versus situational debate whereby for the Health Belief Model, how one perceives the factors is individual but the factors like perceived barriers also involve a situational aspect. Another common issue was the freewill versus determinism debate arguing that the health belief model is freewill because individuals have choice on how they perceive the factors, yet also deterministic as some financial constraints are out of an individual's control. Reductionism versus holism and generalisations were other popular issues covered.

There were a number of weak responses to this question. Some responses wrote about factors leading to non-adherence without clear indication if they were referencing the Health Belief Model or non-adherence. Other weaker responses talked about how the studies were standardised and improved reliability or use of questionnaires provided objectivity without context. A number of responses tried to cover many issues within a paragraph which were often very brief and generic points made. When evaluating in terms of application to everyday life, candidates should be encouraged (practice in advance) to give very specific examples of what could be done. For example, rather than saying 'doctors could be aware that costs might stop people from adhering to medication', make a suggestion about what could be done 'for example, reducing the costs of drugs for chronic/serious conditions (heart medication) by increasing the cost for less serious conditions (acne cream)'.

© 2024

Organisational Psychology

Question 13

Responses varied for this question with some able to give a cognitive limitation or error and provide a contextualised example such as error of omission: it was a demanding project requiring workers to get things done at a faster pace hence overlooking key information. Weaker responses just identified the limitation or error without any example linked to the stem. A number of responses did not receive credit due to not identifying any correct limitation or error.

Question 14

(a) There were some good, full mark responses to this question. These were able to give a clear definition of social loafing. Weaker responses that achieved 1 mark just stated that the group had a negative impact on their performance. Some responses did not receive credit as they described social inhibition instead of social loafing.

(b) Better responses were able to provide a solution e.g. delegating the task across members and explaining how it will reduce social loafing or using electronic performance monitoring. Weaker responses often lacked clarity or simply mentioned how the workers needed to be in a 'team' or 'encourage the team' which was not creditworthy. Some responses were not relevant such as providing closed office spaces or giving a room to each employee which were also not creditworthy.

Question 15

(a)(i) (ii) The stronger full mark responses were able to identify and describe a followership and contextualise it in terms of their factory role. The most common choices were the sheep for **part (a)(i)** and effective follower for **part (a)(ii)**. Weaker responses often did not contextualise their response and apply it to the particular factory role to achieve the second mark. Some did not know this area of the syllabus so could not identify any creditworthy followership style.

(b) Common creditworthy responses were that the style is reductionist as it postulates that people only have one followership style, whereas people may change over time and have a different style. Many weaker responses stated that the 5 followership styles ignored other types of followership but did not explain what these would be. Others stated that followership was a Western concept and not applicable to collectivist cultures (but again, did not explain why this was the case).

Question 16

(a) There was a range of responses to this question covering the full range of the mark bands. There were a minority of strong responses, where candidates clearly described both measures clearly and concisely including the dimensions of job satisfaction. Higher mark responses were able to outline the scoring, examples of categories/questions and items. Weaker responses were not detailed enough with their description of the two measures of job satisfaction or gave incorrect details such as incorrect rating scales.

(b) The marks for this question tended to be between Level 1 and Level 3. Those that demonstrated good knowledge of the measures in **part (a)** achieved higher marks in this question. Most attempted the named issue of psychometrics and some gave strengths and weaknesses with clear examples from **part (a)**. Apart from the named issue, popular choices were individual versus situational, reductionism versus holism, and quantitative/qualitative data. Better responses on the individual versus situational debates talked about how each person perceives the questions and items is personal to them, but factors like salary and security are dependent on situational factors. For the reductionism versus holism debate, responses that achieved in the higher bands discussed how the questionnaires are holistic as it takes into account different aspects of job satisfaction, yet reductionist as well as it limits job satisfaction to quantifiable categories but fails to take into account personal interpretations of the questionnaire and other cultural differences which may contribute to their response. In weaker responses many candidates gave statements which only supported one side of the debate. Weaker responses were not contextualized

Candidates could correctly identify the objective and subjective data in each study and provided examples, e.g., in Fox's study, objective data included the records pertaining to the number of days

© 2024

a worker or his/her team had taken off as a result of an injury and subjective data included the anecdotal data pertaining to the wife driving 50 kms to take advantage of the tokens they had earned before the initiative was going to be stopped. Similarly in Swat's study, the percentage of accidents that occurred in the different categories from the 83 accidents recorded and investigated as well as the subjective data generated from the interviews with the Supervisors and floor managers on their perception of the causes of the accidents. Other evaluative issues that were successfully applied included generalisations, applications to everyday life, longitudinal method (particularly with reference to Fox et al.) and validity (which covered ecological validity, longitudinal method, and the data collection methods).

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

4Uadmission
SINCE 2020

© 2024

PSYCHOLOGY

Paper 9990/32

Specialist Options: Approaches, Issues and Debates 32

Key messages

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

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

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

These questions in this exam asked candidates to evaluate the suggestion (such as the theory or technique) that was outlined in the candidate's response to **part (a)** of the question, for example with a strength or a weakness, or explaining why an interview might be ineffective in the scenario given in the stem of the question. It would be helpful for candidates when doing revision to learn strengths and weaknesses of the theories, techniques, self-reports, treatments, etc. they have learned and put these into their revision notes. They should also practice explaining the evaluation point in the context of the question.

Questions 2, 6, 10 and 14

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

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

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

These questions in this exam came from one or two of the bullet points in the syllabus. This exam asked the candidate to describe two self reports, theories or explanations of a disorder identified in the specification under the appropriate bullet point. For this exam, some of the answers used the incorrect topic area in the syllabus or the description was brief. It could be useful for candidates to create revision notes with the title of each topic area and the description in the bullet point as the header. Alternatively, candidates could create a mind map and put this information in the centre. Some of the responses were too long for the first option (e.g. **Question 4a)**. These type of responses did usually receive a high mark. However, this often meant that the candidate did not have enough time to write a complete response for the second option (e.g. **12a**). It would be useful for candidates to practice writing an appropriate length response to these type of questions. Teachers could set a word limit or do a timed response at home or in class. These should be done without referencing any notes or the textbook while completing the timed response.

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

© 2024

This question will always ask the candidate to evaluate the studies, theories, characteristics/explanations/treatments of disorders or techniques described in **part (a)** of the question. The response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. However, most responses that evaluated using two issues in this exam, achieved in the lower bands due to the response being superficial and often with little analysis. Some responses that considered three issues tended to achieve higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the studies, theories, characteristics/explanations/treatments of disorders or techniques described in **part (a)** of the answer.

The candidate must also provide some form of analysis to access Level 3 and above. This could be done by discussing the strengths and weaknesses of the issue being considered, presenting a counter-argument to the issue under discussion or comparing the issue between two studies and/or theories. The response needs to explain the comparison/strength/weakness or counter-argument with examples from **part (a)** of the question. It was common for responses to state that two theories, for example, were 'similar' or 'in contrast' for an issue without any explanation as to why they could be compared in this way. This is assessed as little analysis. It could be limited analysis if the response gives a number of contextualised examples to support the evaluation points. A conclusion at the end of each issue would be helpful in order to show excellent understanding of the issue under discussion. In order to achieve the requirements of the Level 4 and 5 descriptors, it would be best to structure the response by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure the answer covers this requirement of the question.

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

General comments

The marks achieved by candidates for this session of the 9990 syllabus achieved across the full range of the mark band which was very pleasing to see. Some candidates were well prepared for the exam and showed good knowledge, understanding, application and evaluation throughout their responses. A significant number of candidates were not as well prepared and showed limited knowledge and understanding with brief, superficial and sometimes anecdotal responses. These candidates often had limited evaluation and application skills.

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

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

Comments on specific questions

Clinical Psychology

Question 1

The responses to this question covered the full range of marks. Responses that achieved 3 – 4 marks focused on both genetics and biochemical explanations. Common responses included the dopamine hypothesis linked to the specific symptoms of schizophrenia and sometimes included parts of the brain such as the pre-frontal cortex and basal ganglia. Good responses also included genetic links through twin studies and family studies. Some also linked to the specific genes such as COMT and DISC1. Weaker responses usually just identified the genetic explanation with no further elaboration. In addition, these types of responses often mentioned dopamine with no links to parts of the brain where either under or over activity caused symptoms. There were many generic responses that attempted to explain the genetic link especially

© 2024

to twin studies, though often stated twins were more likely to develop schizophrenia and not the concordance. Some responses described the symptoms of schizophrenia which was not creditworthy.

Question 2

(a) There were some good responses that were able to clearly outline the idiographic versus nomothetic debate.

Common good responses focused on the uniqueness of an individual versus establishing generalisations/laws. Many focused on the methodology of in-depth/qualitative data versus large samples/quantitative data. Weaker, one mark responses showed understanding of both terms but were poorly outlined demonstrating some confusion. Responses that were not creditworthy sometimes outlined qualitative and quantitative data the wrong way around or were crossed out and changed, showing some lack of understanding for a small number of candidates. In addition, some responses confused this debate with other debates such as free will versus determinism and individual and situational explanations.

(b) Many responses were able to achieve one mark for this question and if the strength was put into the context of diagnosing schizophrenia then the second mark could be achieved. Most common strengths given were that idiographic is in-depth and therefore detail of a patient's symptoms, history, etc. can help diagnosis or that the idiographic approach allows for the patient's unique experiences of schizophrenia (e.g. type of delusion) to be discovered. The weaker, one-mark responses did not provide this specific link to the diagnosis of schizophrenia by giving specific symptoms. A significant number of responses explained the strength of treating schizophrenia rather than diagnosing it which was not creditworthy.

Question 3

(a) Responses to this question covered the full range of the mark scheme. Those that achieved 3 – 4 marks suggested two reasons why Leo's mother may think he has OCD in some detail and included obsessions and compulsions linked to Leo with examples such as, time consuming, affecting daily life, checking or ordering objects, believing something bad will happen if compulsion is not carried out. A few suggested the reasons for the disorder e.g. psychodynamic or genetic, and why Leo's mum thought this applied to him. Common suggestions included checking something repeatedly such as if a door is locked, ordering toys and extensive cleaning of bedroom. Fewer candidates were able to give an example of an obsession or suggested a different compulsion. Examples of obsessions included that if the compulsion was not carried out his family will become very unwell or something bad will happen

Weaker responses were varied. Some of these were long generic reasons that attempted to describe features of OCD and Leo's relationship with his parents without the technical language of the diagnostic symptoms. There were others that did not address both obsessions and compulsions so limited the marks they could achieve. Many confused obsessions and compulsions and the example given for an obsession was a compulsive behaviour which was not creditworthy. A significant number of responses suggested that Leo does hand-wash but in secret or only his mum has seen it. This type of response was not creditworthy as it suggests that those with OCD obsessively hand-wash and there are no other types of compulsions.

(b) There were some full mark responses to this question. The most common reasons given for why an interview with Leo may not help to diagnose him with OCD was Leo's embarrassment/not being honest and social desirability or response bias with a specific example from the symptoms of OCD. Weaker responses appeared to misunderstand the question. A significant number of responses gave a weakness of interviews without reference to any specific symptoms. Some gave a weakness of doing research on people with OCD without any reference to interviews which was not creditworthy.

Question 4

(a) Responses varied for this question and covered the full range of the marks available. Level 3 responses were less common with most responses achieving either Level 1 or Level 2. Those that did achieve Level 3 kept their response focused on impulse control disorder. Good responses outlined how positive reinforcement could lead to the development and maintenance of impulse control disorders with some giving a clear example from disorders such as gambling and

© 2024

kleptomania. Another feature of a Level 3 response was to outline schedules of reinforcement and how the disorder of gambling has a variable schedule and how this leads to the maintenance of gambling disorder. In addition, these types of responses also gave a clear outline of Miller's feeling state theory as linked to the disorder(s) including a clear summary of what is meant by feeling state, how this is experienced and why this leads to maintenance.

Weaker responses often lacked detail for one of the explanations or both (more common in Level 1 responses). Many responses gave a brief outline of how positive reinforcement could lead to an impulse control disorder, but this frequently did not give any examples and made it difficult for candidates to receive above L1. This was frequently followed up with an outline of non-creditworthy content such as negative reinforcement, Skinner's study with rats, biochemical explanation of the reward pathway and its effect on dopamine. In addition, a minority of candidates referred to OCD rather than impulse control disorder which also was not creditworthy. Weaker responses for Miller's feeling state theory were often very brief with a mention of intense positive feelings from the impulsive action causing it to be repeated with no reference to what the feelings might be or why this would lead to repetitive behaviour.

(b) The marks for this part of the question did cover the full range of marks available and the most frequent levels awarded were Level 2 and 3. Those that achieved Level 3 and above structured their response issue by issue and often started with the named issue of nature versus nurture, along with clear and specific examples from the explanations from **part (a)** and analysis. Apart from the named issues, other popular issues covered were free-will versus determinism, application to everyday life, and individual versus situational. Popular examples for the named issue included how the positive reinforcement is nurture and why and some made comparisons to the biochemical explanation supporting nature which is not taken into account by behaviourist explanations. Miller's feeling state was often presented as supporting nurture as experiences need to happen for the positive feeling state to emerge and this must be remembered in order to achieve these feelings again. Some responses did explain why it could also be seen to support the nature side of the debate as someone's feelings about the impulsive behaviour could be affected by genetics which in turn affect biochemical responses (e.g. high excitement) and how they respond to situations.

Weaker responses achieving Level 1 or Level 2 did not contextualise their response. While some of the evaluation points were valid, the lack of context prevented responses from achieving a higher band. There were a number of generic definitions and a very weak attempt to include context for a specific impulse control disorder or why the issues apply to this topic area. Many responses evaluated treatments or how to measure impulse control disorder and the parts of the response that included this were not creditworthy unless clearly linked back to explanations which the vast majority did not do.

Consumer Psychology

Question 5

There were several good responses to this question, and some achieved full marks. Full mark responses linked the mobile phone to memory and included proactive interference, retroactive interference and similarity and gave nice examples for the effect on memory of the details of the new mobile phone. Weaker responses did not refer to the types of interference but instead suggested that the consumers would find it difficult to remember the details of the new mobile phone because the adverts were so similar to one another with no reference given to interference at all. These types of responses achieved Level 1. It was also common for responses to refer to the incorrect part of the syllabus such as topic area 2.5 'Advertising' such as product placement and brand awareness which were not creditworthy.

Question 6

(a) Full mark responses showed they had a clear understanding of the reductionism versus holism debate and wrote a good outline of both sides of the debate. Candidates wrote better definitions of holism compared to reductionism with many just stating it is a simple explanation. Weaker responses showed some understanding of both terms but poorly outlined demonstrating some confusion. This type of answer was given 1 mark. Some candidates answered this the wrong way around, with some crossing out, demonstrating some lack of confidence in the debate. 20 per cent of candidates did not attempt to answer this question.

© 2024

(b) Some responses were able to gain full marks. A minority of candidates did define competitor-focused sales technique and then explained why it would be considered to be reductionist. Typical responses included competitor-focused sales techniques ignore product quality and/or needs of the consumer with a focus purely on the competitor's product, advertising strategies, etc. Weaker responses could define competitor-focused sales technique successfully, but did not explain why it was reductionist with a number not even attempting to explain this in their response. There were a few responses that could not define competitor-focused sales techniques, and this often occurred when the candidate merely wrote 'Competitor-focused sales techniques are…'. 19 per cent of candidates did not attempt to answer this question while answering the majority of the other questions in Consumer Psychology.

Question 7

(a) Good 3 – 4 mark responses clearly knew the Engel Kollat Blackwell model of buyer decision-making and linked this to Mariyah's buying a car. Most common responses included problem recognition and alternatives. Some weaker responses just described one stage – searching for alternatives with a comparison made between different types of cars for features such as reliability, cost of repairs, etc. Many of these responses did not identify the stage but instead just stated that it involved making comparisons between different sellers of cars or types of cars to achieve 1 mark.

Some candidates did not know the model so just guessed at what Mariyah might need to do to buy a car such as try to trade in her car, obtain money for the new car or decide what features she should look for in a new car which was not creditworthy. 22 per cent of candidates did not attempt to answer this question.

(b) A few responses were able to achieve 1 or 2 marks for the answer to this question. Common strengths included application to everyday life and a detailed account of the stages. Those that could give an example of how one or all of the stages from the Engel Kollat Blackwell model support their strength could achieve the second mark. It was far more common for responses to achieve 0 marks as candidates did not know the model so therefore could not evaluate it. This did cause a number of candidates to not answer the question, 27 per cent of candidates did not attempt to answer this question.

Question 8

(a) There were some good Level 3 responses to this question. These types of responses included an outline of the pleasure-arousal-dominance (PAD) model with clear links made between the model and the effects on consumers. In addition, Level 3 responses frequently gave an outline of the study on effect of odour by Chebat and Michon. Weaker responses often did not make many or sometimes no links between the PAD model and consumer behaviour. Sometimes an outline was given of generic/anecdotal effects of odour on shoppers such as having the smell of freshly baked bread in the bread aisle of a grocery store. Those that did outline the Chebat and Michon study sometimes gave some incorrect details such as noting down how much consumers were spending/number of products sold when this data was not collected. Many also incorrectly stated that it supported the PAD model which is also incorrect. It was also somewhat common for responses to give detailed descriptions of the study by North et al. on music in restaurants. This study was not investigating the PAD model or odour so was not an appropriate response and no credit was given.

(b) There were a very small number of Level 3 and above responses to this question. Those candidates who had outlined a study in **part (a)** were able to discuss the named issue of qualitative and quantitative data. Some gave several strengths and weaknesses of quantitative data and gave examples from the study by Chebat and Michon. They recognised there was not any qualitative data in the study and the issues that this caused. However, some responses stated there was qualitative data in the study which was incorrect. Other common evaluation points included application to everyday life, cultural differences and strengths/weaknesses of questionnaires.

Most responses for this question achieved in Level 1/Level 2. This was due to writing superficial evaluation points. The issue was named and the model or the study was either stated to support the issue or which side of the debate it supported. This question had a number of very brief responses often with just a few sentences within which numerous issues were identified with little else provided. It was also noted a significant number of candidates did not attempt the question with 27 per cent leaving the answer space blank.

© 2024

Health Psychology

Question 9

There were many strong responses to this question achieving 3 – 4 marks. Good responses outlined two ways Camilla could check with some detail and in the context of whether her father has been taking his pills every day. This was more effectively done when the response outlined the TrackCap with a suggestion of how Camilla could measure that the medication was taken every day. A second way that was often clearly outlined was a urine or blood test. A small minority of responses were able to explain that a home urine kit could be done every day and monitored by Camilla. Many candidates did fail to notice that Camila needs to check every day, so gave responses of checking using a weekly blood test, for example which was awarded 1 mark rather than 2. Weaker responses frequently outlined one way rather than two or two ways without a reference to how Camilla could measure this to check daily medication use. There were a number of suggestions that were not creditworthy such as watching him or giving him a questionnaire to complete.

Question 10

(a) The majority of responses were able to outline what is meant by one practitioner style, but few were able to do this in relation to an individual explanation. The most common choice was the patient centred style. Many candidates confused individual explanations with individual differences and therefore the outline of the style in relation to individual differences was not creditworthy.

(b) There were a few good responses to this question. The most common weakness given was the patient not being listened to so not able to express their viewpoints. Some responses could link this directly to a practitioner style to get full marks. Some responses were able to get one mark through the outline of a weakness without a link to a practitioner style so long as it was a clear weakness of the individual approach. Some responses gave weaknesses of self-reports done by the patient using either patient or practitioner centred style. Examples of this included bias, dishonesty, socially desirable responses, lack of detail or too much detail depending on style chosen. These were not creditworthy responses.

Question 11

(a) There were a number of good responses to this question achieving 3 – 4 marks. Most common reasons included the chest pains with no obvious reason and the multiple abdominal scars. Some included that he was male and some stated that he had had multiple previous hospital visits. Weaker responses often identified the multiple scars and that the doctor could find nothing wrong with Hanif. A one-mark response included one feature of Munchausen shown by Hanif.

Most candidates were able to identify that Hanif appears to be lying about chest pains as far as the doctor can see. Stronger candidates pointed out that previous scars are an essential features of Munchausen syndrome. Some responses outlined the study by Aleem and Ajarim which was not creditworthy, and they did not relate any of their response to Hanif.

(b) Most responses attempted to explain a weakness of the diagnostic features of Munchausen syndrome. Common weaknesses given were the patient not having any physical features to help with diagnosis and the nature of the syndrome means that patients are likely to be pathological liars and so obtaining accurate information on which to base a diagnosis will be difficult. A number of candidates attempted to explain that these features come from a very limited number of case studies on patients because it is such a rare disorder and how this could make the symptoms limited or over exaggerated such as the presence of scars as only a small number of patients with Munchausen have scars. One non creditworthy response that was given by some candidates was that the patient could have a serious illness and with a diagnosis of Munchausen this patient will not get the treatment that they need. This is an issue with diagnosing their actual illness rather than an issue with diagnosing Munchausen.

© 2024

Question 12

(a) There were a small number of candidates who were able to give excellent details of the two questionnaires outlined in the question. This was better for the Holmes and Rahe life events questionnaire than a measure of Type A/B questionnaire. In addition, many candidates received either a Level 1 or Level 2 mark and this was unexpected. A significant number of candidates did not attempt the question (15 per cent). Some were able to outline one or two features of Holmes and Rahe. Often candidates described the studies that were done using these questionnaires. Anything given about the study which was not a feature of the two questionnaires did not receive any credit. This error was often carried forward to **part (b)** where the candidate would evaluate the studies outlined in **part (a).** Other weak responses were where the candidate did not know either questionnaire and appeared to be guessing at the features which occasionally did achieve 1 mark.

There were a small number of responses that achieved Level 3 or above. These types of responses started with the named issue which was questionnaires and were able to give some strengths and weaknesses. Candidate's responses achieved Level 3 and over due to giving clear detailed reference to the questionnaires from **part (a).** The strongest responses also included some analysis. This was rarely well explained and therefore often Level 3 was awarded. If the analysis was explained such as explaining why you are stating that the two questionnaires are similar due to collecting quantitative data, then Level 4 and above could be awarded. Other common issues raised were quantitative data, reliability and objective and subjective data, validity and psychometrics. However, psychometrics was often very similar to the points raised on quantitative data.

There were a significant number of weak responses to this question where many received a Level 1 or 2. This was due to the response doing superficial or basic evaluation. This frequently involved candidates identifying a number of issues and then stating, for example, how each questionnaire has quantitative data. It might give a general strength or weakness for one of the examples but no specific reference will be made to either questionnaire. Context refers to specific examples e.g. stating that Holmes and Rahe collects numerical data is not specific whereas stating that it asks the participants to indicate each life event they have experienced in the past 12 months and that predetermined numerical values from 1 – 100 are given depending on how stressful this life event is, is specific.

Part (a) responses that outlined one or two studies were not given any credit unless it is clearly about the measure of Type A/B. In addition, any evaluation of the explanations of stress rather than the measures was also not given credit. 21 per cent of candidates did not answer the question and left it blank.

Organisational Psychology

Question 13

Responses covered the full range of the mark scheme. Good responses were able link each suggestion to the reduction in conflict between Ulrich's workers which puts the response into context. The most common stereotypically feminine leadership styles used in responses was democratic and/or relationship orientated. Candidates could achieve full marks with one suggestion so long as the suggestion included the correct terminology and gave examples to explain the effect on the conflict between the workers once the change has been implemented.

Weaker responses frequently just identified one or two feminine leadership styles such as relationship oriented without any example linked to the conflict between workers from the stem. There were a number of generic responses that outlined stereotypical feminine behaviour such as softer, weaker, showing empathy without any link to leadership styles which was not creditworthy. Other non-creditworthy responses included to hire more women, and some outlined how Ulrich could be more masculine.

Question 14

(a) There were some good, full mark responses to this question. These were able to give a clear definition of intrinsic motivators at work frequently with examples given such as praise and recognition. Weaker responses stated that intrinsic motivators are motivators that are intrinsic to the person which is not creditworthy as it is tautological. However, these type of responses could

© 2024

achieve 1 mark for an example (or many examples). The most common error was to outline extrinsic rather than intrinsic motivation which received no credit.

(b) The responses to this question varied quite a bit with some giving a clear explanation of why intrinsic motivators can vary between cultures. An example of this is those in individualistic cultures (e.g., USA) are likely to be motivated more by praise (1) than those in collectivist cultures (e.g., China), where praise is immodest (1). Weaker, one-mark responses, attempted to explain how individualistic and collectivist cultures may have different intrinsic motivators (but rarely explained why). Some candidates misunderstood and wrote about extrinsic motivators, despite knowing what intrinsic motivators are in **part (a)**. This type of response was not creditworthy.

Question 15

(a) The marks received by candidates for this question covered the full range of the mark scheme. Those who achieved 3 – 4 marks gave two clear suggestions each linked to one of the levels of the hierarchy and appropriate to the factory environment. Most common responses included safety, belonginess and self-actualisation. Good responses explained the level of the hierarchy and linked it to the factory. For example, using better ventilation to fulfil their safety needs so that they did not inhale dust from the fabric or to fulfil esteem needs by introducing 'factory worker of the week' posted on the wall. The most common levels used for suggestions were esteem, belonging and safety.

Weaker responses could identify which level of the hierarchy Tayyibah could utilise with a mention of how it might improve motivation. However, they were unable to explain what Tayyibah could do in the factory (failing to contextualise). Other weak responses were where a generic explanation of the hierarchy (e.g. being with others or feeling valued) was given with an attempt to explain what Tayyibah could do. Responses that did not receive any credit included referring to Tayyibah using a questionnaire to find out which level of the hierarchy each worker is at which the question does not ask for. In addition, some candidates made links to wages rather than other stages in Maslow's hierarchy of needs that can motivate the workers. The question clearly states that they do receive a good wage and are still not motivated.

(b) There were a number of good, full mark responses to this question. Most common weaknesses included cultural bias and examples of people who have reached the top of the hierarchy without meeting lower needs. Another creditworthy weakness which was frequently done well by those who chose it is that the hierarchy is rigid and inflexible and some people may value one level more than another. For example, they may value love and belonging less than self-esteem even though self-esteem is ahead in the popularity goals. Weaker responses gave a weakness of Maslow's hierarchy without a link to the workplace.

There seems to be some lack of understanding about Maslow's hierarchy as many candidates gave the example of Mother Teresa as not having basic needs met despite reaching self-actualisation. Yet Mother Teresa lived in a closed community where all her basic needs were met so this example was not creditworthy. It is better for candidates to give examples from organisations wherever possible.

Question 16

(a) There was a range of responses to this question covering the full range of the mark bands. There were some good responses outlining both the stages of group development and Belbin's nine teams' roles.

Group development was outlined far more effectively with some clear examples from groups that might form in an organisation compared to Belbin's nine teams' roles which was often list like. Some weaker responses did confuse the stages of group development. Level 1 responses frequently gave a summary of the stages which were sometimes labelled incorrectly or put in the wrong order.

(b) The marks for this question tended to be between Level 1 and Level 3. Those that demonstrated good knowledge of the theories in **part (a)** achieved higher marks in this question. Most attempted the named issue of application to everyday life. Higher mark responses gave specific examples of how the theories can be utilised by organisations. For example, how an organisation could select staff based on what type of role from Belbin's nine team roles they would best suit. Stages of group

© 2024

development can be used to monitor teams as they are formed in an organisation and for management to be aware that conflict within a team is a normal part of its development. Other commonly used evaluation points included generalisability and idiographic versus nomothetic.

Weaker responses lacked evaluation points that provided any depth of explanation. Each point tended to be very brief, and reference was frequently made to the theories in **part (a)** but this was merely identifying which side of the debate the theory was on or whether it did have applications or good generalisability but only occasionally was this explained. Sometimes analysis might be given but this was also just identified and not explained. Frequently this was structured by one paragraph for Belbin and the second paragraph for stages of group development. This led to a lot of repetition and no analysis was given.

4Uadmission
SINCE 2020

© 2024

PSYCHOLOGY

Paper 9990/33

Specialist Options: Approaches, Issues and Debates 33

Key messages

Questions 1, 3(a), 5, 7(a), 9, 11(a), 13, 15(ai), 15(aii) –

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

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

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

Questions 2, 6, 10 and 14 –

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

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

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

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

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

This question will always ask the candidate to evaluate the studies, theories, characteristics/explanations/treatments of disorders or techniques described in **part (a)** of the question. The

© 2024

response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. However, most responses that evaluated using two issues in this exam, achieved in the lower bands due to the response being superficial and often with little analysis. Some responses that considered three issues tended to achieve higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the studies, theories, characteristics/explanations/treatments of disorders or techniques described in the **part (a)** of the answer.

The candidate must also provide some form of analysis to access Level 3 and above. This could be done by discussing the strengths and weaknesses of the issue being considered, presenting a counter-argument to the issue under discussion or comparing the issue between two studies and/or theories. The response needs to explain the comparison/strength/weakness or counter-argument with examples from **part (a)** of the question. It was common for responses to state that two theories, for example, were 'similar' or 'in contrast' for an issue without any explanation as to why they could be compared in this way. This is limited analysis. A conclusion at the end of each issue would be helpful in order to show excellent understanding of the issue under discussion. In order to achieve the requirements of the Level 4 and 5 descriptors, it would be best to structure the response by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure the answer covers this requirement of the question.

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

General comments

The marks achieved by candidates for this section of the 9990 specification achieved across the full range of the mark band which was very pleasing to see. Some candidates were well prepared for the exam and showed good knowledge, understanding, application and evaluation throughout their responses. A significant number of candidates were not as well prepared and showed limited knowledge and understanding with brief, superficial and sometimes anecdotal responses. These candidates often had limited evaluation and application skills.

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

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

Comments on specific questions

Clinical Psychology

Question 1

The responses to this question covered the full range of marks. Responses that achieved 3-4 marks often referred to the process of systematic desensitisation by outlining the steps with accurate terminology. Common suggestions for contextualised responses involved Jude looking first at a picture of food followed by exposing Jude to real food eventually. Many candidates were able to outline the fear hierarchy and muscle relaxation. Weaker responses confused systematic desensitisation with other techniques. It was common for responses to fail to explain how the gradual exposure would occur so they achieved fewer marks overall.

Question 2

(a) There were some good responses that were able to clearly outline the reductionism side of the reductionism versus holism debate. For example, stating that it is explaining psychological phenomena by breaking it down into smaller component parts. One-mark responses often outlined that reductionism is to do with an explanation or concept that focuses on one thing/ignores other

© 2024

factors. Some responses were too specific, using definitions with psychological disorders/diseases as part of the definition which was not creditworthy.

(b) The majority of responses were able to explain the biochemical explanation for schizophrenia through outlining the dopamine hypothesis. However, only a handful were able to explain why this was reductionist (even when they had successfully defined reductionism in **part (a)).** Most attempted to explain why it was reductionist but often just stated it ignored other explanations, or it ignored cognitive explanations which was not creditworthy. Those that did achieve full marks were able to explain in some detail exactly what was ignored by this explanation such as lack of self-monitoring leading to hallucinations.

Question 3

(a) Responses to this question covered the full range of the mark scheme. Those that achieved 3-4 marks often were able to explain 'why' the Doctor would want to use the GAD-7. Common responses included that it is objective, quick and easy, to determine the severity of her anxiety or to see what appropriate treatment can be given. The best responses were able to support their statements with the GAD-7 e.g. quick and easy because it has only 7 items, or if to find out the severity of her anxiety due to her showing one of the symptoms (trouble relaxing) in the GAD-7. Weaker responses lacked any explanation as to why the doctor would want to use the tool. Some responses mentioned just 'quick and easy' without saying why it's quick and easy and achieved 1 mark for this reason given. A number of responses confused the measure (GAD-7) with Generalised Anxiety Disorder itself which was not creditworthy.

(b) There were some full mark responses to this question that could relate the weakness back to Aisha or the anxiety disorder. Those that did not achieve full marks lacked this link. Common weaknesses were that it lacks any in-depth explanation as to finding out the 'why' Aisha has anxiety. Other common responses were that the GAD-7 is a self-report measure which can lead to social desirability or Aisha lying. However, some responses gave incorrect weaknesses which received no credit including how the questions might be confusing and Aisha may not understand the questions. It should be noted the questions on the GAD-7 are very simple to understand. Another weakness that was not creditworthy included that the GAD-7 cannot be generalized to other phobias/disorders.

Question 4

(a) Responses varied for this question and covered the full range of the marks available. Level 3 responses were able to outline the study in detail, covering the participants, procedure (including one or more of the self-reports used to collect data) and results of the study as outlined in the mark scheme. Some responses were able to quote percentages of the results from the study although this was not a requirement for full marks. Weaker responses included fewer details of the study. A minority of responses confused this key study with another, but there were a significant number of responses that simply stated that it was a study comparing CBT face-to-face with that by telephone but with no details at all. Neither of these types of responses were creditworthy.

(b) The marks for this part of the question did cover the full range of marks available with the most frequent levels awarded being Level 2 and 3. Those that achieved Level 3 and above structured their response issue by issue, and often started with the named issue of reliability, along with evidence from the study in **part (a)** and analysis. Apart from the named issues, other popular issues covered were individual versus situational, quantitative data, validity, generalisability and application to everyday life. Popular examples for the named issue included how the timings of the telephone calls were standardised, and that all participants were given the same questionnaire and the sessions in each condition were conducted by the same therapist. Good responses were also able to outline how the study may not have reliability due to the lack of control on the surroundings of the telephone group. Another common response was that each individual's responses were different thus the way that CBT was conducted would not have been the same for every participant.

Common applications to everyday life were that the use of a telephone is less time consuming and can save cost on travel for therapy. Generalisability was also a popular point as participants were from the UK so it's difficult to generalise to other populations. Other responses like individual versus situational mentioned that the study is situational due to the outcome of the therapy sessions being dependent on whether they are in the telephone group or face to face group, yet also individual as each participant gives a unique response during CBT. Other responses covered

© 2024

the strength of using quantitative data from the questionnaires as it allowed for objectivity and comparisons.

Weaker responses achieving Level 1 or Level 2 did not contextualise their response. While some of the evaluation points were valid, the lack of context prevented responses from achieving a higher band. Some who wrote about a debate such as determinism versus free-will did not explain how the study supported the relevant side of the debate. Some provided too many issues with no depth in explaining why.

Consumer Psychology

Question 5

There were several good responses to this question and some achieved full marks. Full mark responses were able to outline two of the effects of changing to freeform layout (engages customer so longer browsing time, more enjoyable to use) and put this into the context of an online grocery retailer. Some outlined how the change might affect the customer such as being confused at the new layout which was also creditworthy. Weaker responses sometimes outlined one effect in context or were too brief (simply stating, for example, that it would be 'more entertaining'). A number of responses described customer behaviour in a physical rather than virtual store which was not creditworthy to achieve full marks.

Question 6

(a) Good responses were able to outline an application of using children e.g. having them watch an advertisement of a product aimed at children versus adults and seeing which brands they recognise. Some outlined showing children a few brands through a game before testing them to see which brand they are likely to remember or outlining the method from the Fischer et al. study. Some were able to relate it back to brand awareness e.g. by doing these it could help researchers gain insight to children's preference which will shape their marketing strategy. Most made some attempt at this question but lacked clarity as to why it would be suitable for children and this was frequently not creditworthy. A significant minority gave a description of the Auty and Lewis study, which is about product placement and not brand awareness so also not creditworthy.

(b) Some responses were able to gain full marks. These types of responses were able to outline a strength e.g. children are in critical stages of cognitive development which reveasl how branding shapes them over time/children are future consumers/children are less likely to have demand characteristics. The strongest responses were able to relate it back to brand awareness e.g. guide marketing for young audiences who will influence their parents' purchases. Weaker responses often lacked clarity and simply stated, for example, it helps understand the impact of advertising on children, without saying how. A response that was not creditworthy was stating that children can be manipulated very easily so companies could exploit this (suggesting unethical methods).

Question 7

(a) There were a number of full mark responses with two clear suggestions of how background music could improve the atmosphere. The most common responses referred to genre of music (almost always classical) and volume. Those that suggested classical music stated it would increase spending and referred to the North et al. study's results. Responses that could not achieve full marks often did not link to how the feature of music would affect the customer/atmosphere.

(b) Stronger, full mark response described how the genre of music needed to match the atmosphere of the restaurant and the effect this would have if this did not happen. Other common responses included that the music may distract customers and make them uncomfortable. Some responses talked about how the volume of music may be too loud and cause customers to leave or that the customers may not like the music/genre of music. There were a few novel suggestions including the cost of installing the music system in the restaurant. Responses that could not achieve full credit often did not link the problem back to the customer. Some described how the loudness of the music may affect the taste of the food, which refers to *noise* (rather than music in research) which was not creditworthy.

© 2024

Question 8

(a) There were many good, Level 3 responses to this question which included an outline Hall's four zones with the inclusion of a study. The vast majority summarised the Robson et al. study on table spacing in a restaurant. Candidates who knew about the overload, arousal and behaviour constraints could provide examples and relate that to restaurant tables. Weaker responses often gave an outline of Hall's four zones but the description of the Robson et al. study lacked detail and many very briefly outlined overload, arousal and behaviour constraint. Some responses did not include overload, arousal and behavioural constraint which limited the mark awarded to Level 2 at best. There were some very weak descriptions for this question with responses showing very little knowledge of the topic area and achieving either Level 1 or 0 marks for this reason.

(b) There were some Level 3 and above responses to this question. Most responses evaluated using the named issue of cultural differences with good examples given about how these differences might affect personal space at tables. Other common issues were application to everyday life, individual versus situational, generalisability (with reference to the Robson et al. study) and reductionism versus holism.

Weaker responses lacked depth in their discussion of cultural differences and simply stated that it could not be generalised outside of Western culture with no examples given which limited the marks awarded to Level 1 or Level 2. There was some confusion in weaker responses about the Robson et al. study with many incorrectly stating that it took place in a restaurant (when it took place as an online survey) so no marks were awarded for the discussion around good ecological validity. Weaker responses also provided a long list of evaluation issues with just a vague connection to the study/theory outlined in **part (a)**.

Health Psychology

Question 9

There were many strong responses to this question achieving 3-4 marks. Good responses made some good, clear suggestions (usually more than one). Strong responses suggested the use of mild fear arousal and described how pictures could be shown of the result of diabetes, obesity and/or heart disease. Others outlined how the teacher could use role models (similar to the Tapper et al. study) and rewards to encourage healthy eating. It was less common for responses to suggest how there could be a reduction in the eating of unhealthy foods and therefore full marks could not be awarded. Weaker responses did not make it clear that their suggestions would be applicable to the teacher's class (the responses could be for a whole school). Other weaker suggestions gave a very brief outline of their idea. Some suggestions were not practical for a teacher in a school or even possible (such as increasing the cost of unhealthy food or changing how it is advertised) and these were not creditworthy.

Question 10

(a) The majority of responses were able to outline what is meant by freewill. An example of a strong response is 'freewill is the individual's own autonomy and ability to choose to behave as they desire. For example, positive psychology states that individuals who have a happy life will choose to focus on their virtues. As they can choose this focus on their virtues, they can choose to be happy or decide against it as per their own accord.' Weaker responses did not provide an example from positive psychology. Some examples simply just said to 'think positively' as an example which was not creditworthy. Some responses gave a tautological definition which could not achieve credit.

(b) There were a few good responses to this question. Responses that achieved credit for this question often used examples from the Shoshani and Steinmetz study mentioning how candidates from single-parent / low-income families find it more difficult to achieve one of the lives due to their situation being determined by socioeconomic factors. Or the response used an example from meaningful life such as being able to do charitable work required the person to have access to this in their personal life which was not in their control. Many responses were not creditworthy due to lack of any reference to positive psychology and just stated, for example, that some people would never have good mental health (which clearly goes against the premise of positive psychology). Many responses gave a reductionist explanation instead of a deterministic one which also was not creditworthy.

© 2024

Question 11

(a) There were many good responses to this question. Most were able to outline at least one biological measure. Common responses were salivary cortisol, measuring heart rate using a pulse oximeter or measuring blood pressure. Many could achieve full marks for one or more of the biological measures by explaining how it would be used to make comparisons between home and work and/or what a high or low result on the measure meant in terms of stress. Lower mark responses just identified the measure. A significant number of responses suggested the use of an fMRI. This would clearly be unsuitable because, although such a device could be used as a biological measure of stress, fMRI machines are not portable devices (and could not be present in Nadim's work and home). A minority suggested non-biological measures of stress such as semi-structured interviews which was not creditworthy.

(b) There were some full mark responses to this question that explained the weakness of one of the measures from **part (a)** and put this into the context of measuring stress. The most common responses were that the measure may not truly reflect Nadim's stress levels and that his stress levels could be caused due to other named examples. Some candidates who talked about taking blood cortisol levels mentioned it will be troublesome for Nadim to take a blood sample at his office. No credit was given if the response explained that the test might not be administered correctly. If candidates evaluated the use of an fMRI (due to its unsuitability in **part (a)**), this was not creditworthy either.

Question 12

(a) The responses to this question covered the full range of the mark scheme. Better responses gave clear and often detailed and accurate descriptions of the Health Belief Model and reasons for non-adherence, with the strongest frequently giving details of the Laba et al. study. Weaker responses were able to achieve marks from writing about their knowledge of the Health Belief Model. Fewer responses were able to give a detailed outline of the rational non-adherence. Some only identified factors of the Health Belief Model and did not elaborate further on how these were barriers. Some responses were brief with short outlines of the Laba et al study. There were a significant number of candidates who did not know any content about rational non-adherence. There were also several responses which outlined the effect of doctor's clothing and tone of voice which was not creditworthy.

(b) The marks for this question covered the full range of marks. Most responses attempted the named issue of applications to everyday life. Common applications included using knowledge of the health belief model to lower medicine cost, do promotional campaigns on the benefits of medicine, and educating the public on the severity of certain illnesses. Another common issue was the individual versus situational debate whereby for the Health Belief Model, how one perceives the factors is individual but the factors like perceived barriers also involve a situational aspect. Another common issue was the freewill versus determinism debate arguing that the health belief model is freewill because individuals have choice on how they perceive the factors, yet also deterministic as some financial constraints are out of an individual's control. Reductionism versus holism and generalisations were other popular issues covered.

There were a number of weak responses to this question. Some responses wrote about factors leading to non-adherence without clear indication if they were referencing the Health Belief Model or non-adherence. Other weaker responses talked about how the studies were standardised and improved reliability or use of questionnaires provided objectivity without context. A number of responses tried to cover many issues within a paragraph which were often very brief and generic points made. When evaluating in terms of application to everyday life, candidates should be encouraged (practice in advance) to give very specific examples of what could be done. For example, rather than saying 'doctors could be aware that costs might stop people from adhering to medication', make a suggestion about what could be done 'for example, reducing the costs of drugs for chronic/serious conditions (heart medication) by increasing the cost for less serious conditions (acne cream)'.

© 2024

Organisational Psychology

Question 13

Responses varied for this question with some able to give a cognitive limitation or error and provide a contextualised example such as error of omission: it was a demanding project requiring workers to get things done at a faster pace hence overlooking key information. Weaker responses just identified the limitation or error without any example linked to the stem. A number of responses did not receive credit due to not identifying any correct limitation or error.

Question 14

(a) There were some good, full mark responses to this question. These were able to give a clear definition of social loafing. Weaker responses that achieved 1 mark just stated that the group had a negative impact on their performance. Some responses did not receive credit as they described social inhibition instead of social loafing.

(b) Better responses were able to provide a solution e.g. delegating the task across members and explaining how it will reduce social loafing or using electronic performance monitoring. Weaker responses often lacked clarity or simply mentioned how the workers needed to be in a 'team' or 'encourage the team' which was not creditworthy. Some responses were not relevant such as providing closed office spaces or giving a room to each employee which were also not creditworthy.

Question 15

(a)(i) (ii) The stronger full mark responses were able to identify and describe a followership and contextualise it in terms of their factory role. The most common choices were the sheep for **part (a)(i)** and effective follower for **part (a)(ii)**. Weaker responses often did not contextualise their response and apply it to the particular factory role to achieve the second mark. Some did not know this area of the syllabus so could not identify any creditworthy followership style.

(b) Common creditworthy responses were that the style is reductionist as it postulates that people only have one followership style, whereas people may change over time and have a different style. Many weaker responses stated that the 5 followership styles ignored other types of followership but did not explain what these would be. Others stated that followership was a Western concept and not applicable to collectivist cultures (but again, did not explain why this was the case).

Question 16

(a) There was a range of responses to this question covering the full range of the mark bands. There were a minority of strong responses, where candidates clearly described both measures clearly and concisely including the dimensions of job satisfaction. Higher mark responses were able to outline the scoring, examples of categories/questions and items. Weaker responses were not detailed enough with their description of the two measures of job satisfaction or gave incorrect details such as incorrect rating scales.

(b) The marks for this question tended to be between Level 1 and Level 3. Those that demonstrated good knowledge of the measures in **part (a)** achieved higher marks in this question. Most attempted the named issue of psychometrics and some gave strengths and weaknesses with clear examples from **part (a)**. Apart from the named issue, popular choices were individual versus situational, reductionism versus holism, and quantitative/qualitative data. Better responses on the individual versus situational debates talked about how each person perceives the questions and items is personal to them, but factors like salary and security are dependent on situational factors. For the reductionism versus holism debate, responses that achieved in the higher bands discussed how the questionnaires are holistic as it takes into account different aspects of job satisfaction, yet reductionist as well as it limits job satisfaction to quantifiable categories but fails to take into account personal interpretations of the questionnaire and other cultural differences which may contribute to their response. In weaker responses many candidates gave statements which only supported one side of the debate. Weaker responses were not contextualized

Candidates could correctly identify the objective and subjective data in each study and provided examples, e.g., in Fox's study, objective data included the records pertaining to the number of days

© 2024

a worker or his/her team had taken off as a result of an injury and subjective data included the anecdotal data pertaining to the wife driving 50 kms to take advantage of the tokens they had earned before the initiative was going to be stopped. Similarly in Swat's study, the percentage of accidents that occurred in the different categories from the 83 accidents recorded and investigated as well as the subjective data generated from the interviews with the Supervisors and floor managers on their perception of the causes of the accidents. Other evaluative issues that were successfully applied included generalisations, applications to everyday life, longitudinal method (particularly with reference to Fox et al.) and validity (which covered ecological validity, longitudinal method, and the data collection methods).

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

4Uadmission
SINCE 2020

© 2024

PSYCHOLOGY

Paper 9990/41

Specialist Options: Application and Research Methods 41

Key messages

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

General comments

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

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

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

Section A

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

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

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

Section B

Candidates should only answer one question from this Section.

4Uadmission

© 2024

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

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

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

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

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

Answers to **part (a)** questions in this section should include an appropriate plan, have applied a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Many answers listed features such as 'I would have a random sample' and 'It would be an independent measures design' without explanation of why it would be a random sample, or how it would be obtained. Elaboration of these general sentences should be included.

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

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

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

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

Comments specific to questions

Section A

Question 1

(a) (i) One mark was awarded for candidates stating that a clinical interview is an interaction between a medical practitioner and a person/patient. Additional marks were awarded for stating that this would take place in a medical setting, where verbal and non-verbal exchange took place and the purpose of which was to diagnose, assess and treat. Some candidates incorrectly assumed that a clinical interview was part of a 'study' with a researcher asking questions.

© 2024

(ii) Participants were interviewed by a second psychiatrist (as the question stated) mainly to confirm the data gathered in the first interview. Many candidates understood this and were awarded full marks. However, some candidates wrote as if this was the first interview to gather initial data and so could not be awarded full marks.

(b) Nearly all candidates wrote about researcher bias, or lack of consistency between the two psychiatrists and were awarded one mark. Some candidates correctly applied their point to the key study and were awarded a further mark, but some candidates did not do this and could only be awarded one mark.

(c) Most candidates had little difficulty in providing a strength and a weakness but frequently failed to locate the strength or weakness in the context of the study. For example, a commonly stated weakness was 'participants might not want to answer because of social desirability'. This is far too general, and it needs an explanation to be awarded marks. Writing in addition, for example, 'because they are embarrassed to acknowledge their family history of depression' makes a good answer worth two marks.

Question 2

(a) Some answers were vague with no more than 'the urge to set fires' but were awarded one mark. Candidates adding more detail, such as 'persistent fascination or preoccupation', 'experience of pleasure, excitement, relief or gratification during or after fire setting' were awarded full marks.

(b) Some candidates suggested giving pyromaniacs questionnaires or interviewing them, but as both these are self-reports no marks were awarded (the question stated …*other than by self-report)*. Answers scoring full marks suggested conducting a covert observation (for example by a family member) of a pyromaniac with a checklist of typical behaviours and recording the frequency of different fire-related activities.

(c) There were some excellent strengths included such as a self-report giving the tester a measure of the extent of the problem or that a self-report measure can help to understand the characteristics of pyromania such as its addictive nature. However, whilst some candidates related these points to pyromania and were awarded full marks, others did not and so marks were restricted. Full marks are awarded for a strength (one mark) plus applying it to the sub-topic (+one mark).

Question 3

(a) (i) Any two features of the sample were awarded marks, which included: 180 customers, 118 female 62 male, aged 16 – 80 years, lived in/around Lund, Sweden. Most candidates were awarded two marks. Some candidates wrote incorrect numbers, some described the supermarket (not a feature of the participants), and some believed the study was conducted in Germany, probably confusing this with the Becker et al. key study.

(ii) Correct identification of the sampling technique, opportunity sampling, achieved one mark and elaboration of this, such as 'approached as they passed the tasting stall' or 'passed the stand in the supermarket' was awarded an additional mark. Many candidates did not know the answer because they guessed at 'volunteer sample' or 'random sample'.

(b) Correct suggestions could only be 'volunteer sample' or 'random sample', which most candidates did and by stating how these techniques could gather participants, were awarded full marks. Some candidates made a correct suggestion and with no elaboration were awarded one mark. Some candidates incorrectly suggested an opportunity sample and could not be awarded any marks.

(c) Like other **part (c)** questions, candidates gave a strength and weakness without relating it to the question, and sometimes answers were very vague. For example, 'there is no selection bias by researchers'. Answers like this receive no marks because there is not even an indication as to what sampling method is being commented on and Examiners cannot make assumptions for candidates. An strength awarded full marks might be '…because the participants are chosen randomly, there is no selection bias by the researchers in their study of tasting jam and tea in a supermarket.'

Question 4

© 2024

(a) A few candidates gave tautological answers which simply re-wrote the question such as, 'a customer-focused sales technique focuses on the customer' and these received no credit. Most candidates provided very good answers which showed good understanding of the concept. Such answers mentioned considering the needs of the customer, letting the customer talk and gathering information about customer preferences and the uniqueness of each customer.

(b) One effect of a customer-focused' sales technique is that the customer feels valued, that they were listened to, and that they are happy. This means that they are likely to return to buy another new car and probably recommend the dealer to their friends. Answers like this were provided by most candidates who were awarded full marks.

(c) Many answers included a strength and a weakness of competitor-focused techniques. A number of candidates linked their points to buying a car, following on from **(b)**, and others chose different examples. Overall understanding of these techniques was very good, perhaps because candidates had direct knowledge of how sales techniques are applied.

Question 5

(a) (i) This question asked *how* participants were recruited, not to identify the sampling technique. As the study states: 'a random number generator was used…' and any elaboration of this, such as 'four patients from each surgery', 'from an inner London general practice', or 'over a four-month period' were awarded the additional mark because they showed knowledge of the key study. Candidates stating nothing more than 'it was a random sample', or any other sampling technique scored no marks.

(ii) This question part was answered correctly by nearly all candidates who knew that 'face-down cards were turned over by the doctor as the patient entered the room' to determine whether the directing or sharing style was used.

(b) Many candidates were able to suggest that not randomly allocating patients to conditions could lead to bias and so questioning the validity the study. Often included was the comment that doctors might select a style for a patient because they knew the patient, or the patient 'looked' like a 'sharing' sort of person.

(c) Some candidates showed good insight by realising that random allocation prevents both participant bias and doctor (in this instance) bias. This, when linked to practitioner style and patient satisfaction could achieve full marks.

Question 6

(a) Attention diversion to manage pain was clearly explained by most candidates who often suggested that it could be passive or active (such as doing a puzzle). A few candidates confused it with non-pain imagery or cognitive redefinition.

(b) There are two alternative (as defined by the syllabus) treatments for pain which are acupuncture and stimulation therapy (such as TENS). Some candidates only identified one of these for 1 mark, but others provided an explanation and/or example for 2 marks. A few candidates suggested taking medicine/drugs but this is not a syllabus defined alternative and could not be awarded marks.

(c) There were some excellent answers with strengths focusing on the ease of use for patients and weaknesses focusing on the difficulty of concentrating enough to divert attention, or that the technique applied more to mild pain and not to chronic pain.

Question 7

(a) (i) In order to be awarded marks, answers to this question needed to show some knowledge of the Swat key study. Swat defined accident frequency as 'the number of *injury* accidents with sick leave days off work per 100 employees in a year. Candidates rewording the question and writing 'accident frequency is the number of accidents' scored no marks.

(ii) To be awarded marks for this question candidates needed to identify the two industrial plants in the Swat study which had the highest accident frequency rate. The two were 'foundry' (5.9) and 'meat

© 2024

processing' (2.8). Any other industrial plant, such as 'machinery' (2.1), or 'furniture' (2.3) could not be awarded marks.

(b) Many candidates correctly wrote about a psychological technique specifically to reduce accidents, such as using a token economy system as outlined by Fox et al (1987). Others applied their more general psychological knowledge and wrote about positive reinforcement if correct safety procedures are followed. A third type of answer often scored no marks with comments such as 'tell them off if they make a mistake'.

(c) The most common strength was that 'it is quantitative and so can be compared' (which is correct for one mark) but if there was no elaboration beyond this and no mention at all about accidents then no further marks could be awarded. The common weakness about accident frequency was that many accidents go unreported or that only major accidents are recorded, not minor ones.

Question 8

(a) Nearly all candidates could outline Maslow's 'physiological needs' and 'social needs' with little ambiguity. Some candidates described the other needs and even the hierarchy even though no additional marks were available for doing this.

(b) Candidates frequently suggested 'doing a questionnaire' for which no marks could be awarded because this is far too vague. The question technique (face-to-face, postal, online) and/or the questionnaire format (open and/or closed questions) should always be included. Candidates being awarded full marks suggested doing a questionnaire with open questions, which would allow the person to talk about their achievement of social needs with colleagues in the workplace.

(c) Only a few candidates demonstrated good knowledge of telephone interviews because most answers were anecdotal, and so their marks were restricted. One strength is that a potentially diverse sample from anywhere in the world could be interviewed, and a weakness is that there is no non-verbal communication, which is often a useful indicator in an interview. Very few candidates related this to need theories of motivation.

Section B

Question 9

(a) Many candidates planned an experiment and the IV was usually MAOIs compared to SSRIs. Some candidates vaguely had 'antidepressants and controls' being unable to name the anti-depressants. The design was commonly an independent design, but some candidates incorrectly claimed repeated measures because each participant repeated taking the same drug. Some candidates had a long list of appropriate controls, such as taking no other drug during the trial and taking the drug for 6 months. Some candidates included a control group with no medication, but this would not allow determination of which antidepressant was more effective. DVs were often not operationalised although some good answers used the Beck Depression Inventory (BDI) to measure the level of depression before and after the trial. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) For psychological knowledge many candidates described the way in which both SSRIs and MOAIs work. Some described the characteristics of depression, and some described the features of the BDI.

(ii) A number of candidates linked what they had written in **(b)(i)** with what they had planned in **(a)**. For example, they might describe SSRIs and link that to why in their plan participants would take one tablet daily or why they conducted the study over a six month period as they needed to allow time for the medication to work. However, some candidates incorrectly described an aspect of methodology (with no link to **(a)**), some continued what they were describing in **(b)(i)** and some candidates could not provide an answer.

(c) (i) Most candidates answered this question part correctly by giving a reason why they used an independent design. Those using repeated measures in their plan struggled to give a reason for its

© 2024

use. A few candidates incorrectly thought 'experimental design' meant 'type of experiment' and explained why they had conducted a laboratory or field experiment.

(ii) The most common response was that an independent design requires more participants. This is not necessarily the case; the design does not determine the number of participants: if the sample size is ten participants in each group, then ten it is. Further, there is the counter argument that more participants might be needed but for less time; a repeated design might have fewer participants, but for double the time. Those using the 'more participants' argument often failed to relate it to their plan.

(iii) The choice of type of experiment produced interesting responses. Some planned a field experiment because participants would be taking the drug in their own home; others planned a laboratory experiment because participants would go to the laboratory once a day to receive their medication. Either argument was legitimate if a reason was provided and full marks could be awarded if it was related to the plan.

Question 10

(a) This question required use of a questionnaire with closed questions. Most answers were good but often did not explain in enough detail for full marks to be awarded. For example, writing 'I chose an opportunity sample' without a comment on how this would be done. Similarly, 'the questionnaire has a four-point scale' without saying what the four points are. Details of the questionnaire technique were commonly absent with nothing more than 'I then gave participants the questionnaire' rather than explaining why they had chosen a particular technique. A simple 'I gave participants a questionnaire and a pencil so they could fill it out before they left the supermarket' would be sufficient. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) For psychological knowledge most candidates could explain multiple unit selling although sometimes examples given in support were ambiguous. The study by Wansink et al. (1998) was also described with some candidates quoting the Wansink examples and data (such as 45 per cent more sales when using multiple unit pricing).

(ii) This question part required an explanation to show how what was described in 10**(b)(i)** informed the plan in **part (a)**. The best format is: **part (a)** should include a specific example of multiple unit pricing that can be used on participants. **Part (b)(i)** should describe what multiple unit pricing is and an example of it (perhaps from Wansink et al). Question **part (b)(ii)** then explains the reasoning behind the example used in **(a)** because it is based on the knowledge described in **(b)(i)**.

(c) (i) Question technique concerns how the questionnaires are administered to participants which could be online, postal or pencil/paper (face-to-face); questionnaire format is about whether questions are open and/or closed. Candidates opting to give a reason for the latter were awarded no marks. Any of the three techniques were appropriate and to answer questions participants could do the questionnaire on-line at home; they did not need to be in a supermarket before answering the questions.

(ii) Many candidates stated nothing more than 'the reason why is not known' which is far too vague because this could apply to anything. For marks to be awarded there must be some comment about the study that has been planned. For example, 'the reason why people opt to choose multiple units (e.g. 3 for $30) rather than a single one (1 for $10) will not be known.

(iii) Many candidates could identify an appropriate descriptive statistic. What they could often not do was to give a reason why they chose that particular statistic. For example, a candidate might write 'I chose the mean to analyse the data' without giving the reason why the mean was chosen over the median or mode. Answers often failed to be related to the plan.

Question 11

(a) Many candidates decided to use every aspect of the McKinstry and Wang study and often their answers were merely a description of that study. Some candidates even included the results found

© 2024

by McKinstry and Wang. Candidates should plan their own investigation that may use one or two aspects of a published study. For example, photographs could be used but the photographs should not be exactly the same as the original. Also, real people could be used rather than photographs. Candidates must show that they can think for themselves and plan their own individual investigation. Some candidates did plan something original, and these answers were interesting to read. Another weakness in answers was that candidates decided not to use a face-to-face interview, instead using a completely different method. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) For psychological knowledge the most appropriate research was that by McKinstry and Wang (1991) who compared five styles of a doctor's dress, including a white coat. Notably that is an 'e.g.' study meaning that an alternative could be used. However, candidates either described the McKinstry and Wang study or no study at all. No candidate described an alternative study.

(ii) As applies to all **(b)(ii)** questions candidates should explain how they used the knowledge described in **(b)(i)** to inform their plan. Some candidates did this very well and were often awarded high marks. Some candidates did not answer the question correctly and could not be awarded marks.

(c) (i) Candidates are advised to read all parts of the question before starting their **part (a)** answer, paying particular attention to the bullet points, because those points will be asked about in **part (c)** questions. In this instance candidates got to this question realising that they had not addressed how the data their plan gathered would be interpreted or scored. This meant that answers were often vague and not linked to the plan.

(ii) Knowledge about face-to-face interviews is very general given the answers to this question. This method should be given just as much attention as any other and so candidates should know at least two strengths and weaknesses. Comparisons with other types of interviews such as telephone should also be known. Most candidates simply stated that 'participants might lie to the interviewer' which might be the case, but they are perhaps more likely to tell the truth because they are face-to-face.

(iii) Many candidates are still confused about the difference between interview *technique* and interview *format*. In this instance the question asked about interview format which refers to structured, semi-structured, or unstructured interviews. Again, candidates should have paid attention to the 'interview format' bullet point and should have used a specific format (structured, semi-structured, or unstructured) in their plan, rather than, for example 'I would ask five questions'.

Question 12

(a) Most candidates planned an experiment, and many applied a number of features specific to experiments, such as IV, DV, controls, experimental design and type of experiment. An interesting range of physical work conditions was included, such as seat design in office workers or the colour or type of lighting (LED versus halogen), although some candidates still used temperature as an IV despite the question telling them not to. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) The syllabus states 'impact of physical working conditions on productivity' and so candidates should have relevant psychological knowledge of several conditions, most logically lighting, temperature and noise depending on the industry, but also office design (e.g. open plan) or other features could be included. The study at the Hawthorne plant could also be included, best described by Kompier (2006).

(ii) There were two types of answer: candidates who explained and related what they had planned in **(a)** to what they had described in **(b)(i)** and so could be awarded full marks and candidates who did not answer the question at all, or those who provided a partial answer.

© 2024

(c) (i) This asked about the choice of DV and some candidates merely referred to 'productivity'. Any DV must be fully operationalised which was a 'free choice', given there was no specific industry identified in the question. A DV could have been 'the number of cuddly toys produced each hour'.

(ii) The choice of type of experiment was often a field experiment, conducted in some factory, which is logical given the nature of the question. Less logical were candidates who planned a laboratory experiment which is a more controlled, but a less realistic situation. Importantly, marks were awarded for the reason given, whichever type is used, and if the reason is appropriate and linked to the plan then full marks can be achieved.

(iii) One or more ethical guidelines must be *included in every plan whatever the question* and so giving a reason for one of these guidelines should have been straightforward. However, commonly there was nothing more than 'I debriefed participants', which is not a reason *why* this is done, and such a brief comment could relate to anything rather than the specific plan of the study.

4Uadmission
SINCE 2020

© 2024

PSYCHOLOGY

Paper 9990/42

Specialist Options: Application and Research Methods 42

Key messages

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

General comments

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

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

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

Section A

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

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

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

Section B

Candidates should only answer one question from this Section.

4Uadmission

© 2024

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

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

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

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

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

Answers to **part (a)** questions in this section should include an appropriate plan, have applied a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Many answers listed features such as 'I would have a random sample' and 'It would be an independent measures design' without explanation of why it would be a random sample, or how it would be obtained. Elaboration of these general sentences should be included.

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

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

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

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

Comments specific to questions

Section A

Question 1

(a) (i) Most candidates could be awarded one mark for stating that an exposure hierarchy is 'a list of fears that are ranked according to the amount of anxiety that each fear creates'. Some candidates went further by giving an example, typically related to the question on blood phobia, or provided a more detailed explanation of the term. A few candidates incorrectly focused on exposure therapy.

(ii) This question required two examples of the exposure hierarchy used in the key study by Chapman and DeLapp. Their list had ten examples, such as 'getting a phlebotomy', 'taking blood pressure

© 2024

myself' and any two from these ten were creditworthy. Answers not on this list could not be awarded marks. Some candidates provided logical but incorrect answers based on the study, and others provided examples about button or dog phobias. The question included the words '… in this study' which candidates should never ignore.

(b) A common suggestion was to use a clinical interview whereby a practitioner could ask the person in detail about their phobia. Also common was the suggestion of conducting some form of observation, with the best answers suggesting an overt observation in a controlled setting where the clinician could control the situation.

(c) Most candidates had little difficulty in providing a strength and a weakness of a case study, such as 'information in depth' and 'inability to generalise'. What many candidates did not do was address the essential part of the question 'as used in this study', requiring candidates to apply the strength and weakness to 'T', the participant in the study. For example, writing that "'T' was a 42-year-old male" would support the weakness of an inability to generalise and show knowledge of the key study.

Question 2

(a) Most candidates were awarded 0 marks because they did not answer the question set. The question asked 'how a person's level of depression is **interpreted** using the BDI'. This meant that describing the BDI, such as examples of questions, was incorrect. Interpretation referred to adding up the number of 0/1/2/3 allocated to each of the 21 questions giving a minimum score of 0, and maximum 63 (21 × 3) with the scores out of 63 applied to 6 levels of depression e.g. 0 – 10 normal; through to over 40 being extreme depression.

(b) Candidates were required to suggest how depression could be measured. Many candidates opted for a clinical interview, which was appropriate but often failed to explain what *measure* would be used. Some candidates suggested using a structured observation with the number of depressive behaviours being measured, but often not stating what those behaviours might be. Answers achieving full marks opted to use a questionnaire and this was acceptable provided it did not replicate any feature of the BDI.

(c) Many candidates incorrectly wrote about social desirability which is a weakness of a participant doing any study/questionnaire; it is not specifically a weakness of the BDI. Many candidates who knew the BDI wrote correctly about: the ambiguous wording of some questions; that 21 questions might be too brief to cover the range of depressive symptoms; the assumption that every person's depression can be assessed by these 21 questions; that the range of 0/1/2/3 is too narrow a scale; that the data gathered is quantitative, ignoring reasons for choosing a specific number. All these correct answers showed knowledge and understanding of the BDI.

Question 3

(a) The question asked about the two parts of the online questionnaire. Some candidates appeared not to know that the study involved a questionnaire; others guessed that it was about personal space. Those being awarded full marks wrote the following: **Part 1**: general participant characteristics (1 mark): and one from: age, gender, ethnicity, place of residence, restaurant use frequency and whether they have restaurant work experience (+1 mark). **Part 2**: measures of emotional, intentional, and anticipated behavioral reactions (1 mark) to one of three images of tables for two placed at a distances of 6, 12, or 24 inches away from each other (+1 mark).

(b) The question required an observation to be used and this provided an opportunity for candidates to apply their knowledge of observations such as whether it would be covert, non-participant, controlled, or structured (the most logical features to use) and perhaps with two observers using behavioural categories. Many candidates did explain the observation in some detail and when applied to a way to assess personal space full marks were awarded. Other candidates wrote nothing more than 'I would observe people in a restaurant' which was too vague to be awarded marks.

(c) Like other **part (c)** questions, candidates often gave a strength and weakness without relating it to the question. Those addressing the question specifically, were often awarded full marks, often stating that using an online questionnaire would lack ecological validity (1 mark) because they were

© 2024

not physically present in a restaurant at a 6' table distance (+1 mark). All answers in question **part (c)** for all options require that the strength or weakness be applied to the study in question.

Question 4

(a) A small number of candidates thought an EEG (electroencephalogram) was a device involved in eye tracking which could not be awarded any marks. For the candidates who knew the study (or even what an EEG measures) they could explain that if some form of advertising media was presented, the increase or decrease in cortical activation could be measured.

(b) Suggestions included the conducting of experiments involving eye-tracking (correct here), questionnaires, interviews and observations (with or without cctv). Answers being awarded full marks explained how their investigation would be conducted and referred to some form of advertising media (such as television, printed or internet).

(c) Many answers gave a strength and a weakness of conducting a study using EEG and some even related their answer to advertising media. Strengths were often about the objectivity of 'brain waves' and weaknesses were commonly the uncomfortable nature of wearing the equipment or the lack of reality in wearing the headset when looking at advertising media.

Question 5

(a) (i) Most candidates were awarded 0 marks when writing 'it measured self-esteem on a four-point scale' because this is nothing more than re-writing the question. What received credit was any two of the following five features from the opening sentences of the original study: 'a ten item'; 'evaluation of worthiness as a human being'; a 'Likert-type scale'; 'scale of 1 – 4'; 'strongly disagree to strongly agree'.

(ii) This question focused on another scale used by Shoshani and Steinmetz, and again most candidates merely re-wrote the question. For this scale there were things in common with those for **part (a)(i)** such as it being a 'Likert-type' and 'ten items' but differed in relation to: '1 – 4 from not at all through to exactly true' and it was 'to predict coping with daily hassles and stressful events'. Credit was also awarded for an example of a question.

(b) Many candidates suggested an appropriate strategy such as providing information or fear arousal. The use of positive reinforcement was also credited. Those relating their suggestion to positive mental health were awarded full marks. Some candidates 'forgot' that the question was about mental health and wrote incorrectly about the 'Food Dudes' or dental hygiene.

(c) The question specifically asked about a strength and weakness of 4-point rating scales. Answers that were about social desirability could not be awarded marks because social desirability is about the person, not specifically a weakness of a 4-point scale. Correct answers focused specifically on a 4-point scale, sometimes contrasting the strength and weakness when compared to a 5-point scale. Answers addressing the *in this study* component of the question could achieve full marks.

Question 6

(a) One mark was awarded for candidates stating that a clinical interview is between a medical practitioner and a person/patient. An additional mark was awarded for stating that this would take place in a medical setting, or where verbal and non-verbal exchange took place or that the purpose of which was to diagnose, assess and treat. Some candidates incorrectly assumed that a clinical interview was part of a 'study' with a *researcher* asking questions.

(b) The question required an observation be used and like **Question 3(b)** this provided an opportunity for candidates to apply their knowledge of observations. One or more features of an observation were creditworthy (e.g. overt observation by a medical practitioner) along with an example to show pain such as facial expression of distress, or distorted ambulation (rubbing or holding). Applying the whole of the UAB was incorrect (it is for patients in hospital for a period of time), but using appropriate parts of it was acceptable.

(c) Most candidates were awarded one mark by stating that a clinical interview does not *measure* anything specifically. Candidates relating this to measuring pain were rewarded with full marks.

© 2024

Another weakness is that a patient may not be able to describe their pain and so the practitioners 'measure' of pain might not be accurate.

Question 7

(a) (i) The sample of participants in the Giacalone and Rosenfeld key study was divided by a median split of 25.5 (1 mark) on the responses to the Sabotage Methods Questionnaire. Those above 25.5 were 'high reason accepters' and those below were 'low reason accepters' (+1 mark) which showed correct knowledge of the study. Candidates with guesses such as 'were randomly allocated' achieved no marks.

(ii) There were many incorrect guesses often suggesting that participants had either committed an act of sabotage or not. The main differences were that high-reason accepters were more likely (than low-reason accepters) to justify: production slow-downs, destruction of machinery/premises/products, and causing chaos.

(b) Many candidates are still confused about how reliability can be tested for a questionnaire. There are two ways: split half and test-retest. Test re-test is where the *same* participants are given the *same* questionnaire, in this instance the Sabotage Methods Questionnaire, at a later time. Inter-rater reliability applies to observations (two independent observers independently observing the same thing) and so not relevant here.

(c) Candidates frequently provided generalised answers about volunteers and although these were creditworthy for 2 marks (out of 4) for two weaknesses, they were rarely applied to sabotage as the question required.

Question 8

(a) A few candidates incorrectly suggested that an open plan office is working outdoors and some suggested that it is working from home. Neither of these answers is correct. An open plan is where there are: no interiors walls or full ceiling partitions; no use of small, enclosed rooms such as private offices. There are large, open spaces; no private place. Candidates outlining any two of these features were awarded full marks.

(b) Two marks were awarded to answers which gave reasons why concentration might be affected which could include: background noise levels, other people talking, people walking about, lack of privacy due to low partitions. Credit was awarded for one suggestion in detail, or two different reasons identified.

(c) Nearly all candidates could give two reasons why generalisations could not be made such as 'only one study', 'only done in one country' and similar. Many candidates struggled to give an example from a study on open plan offices, this could be from Oldham and Brass (1979) or an alternative. Candidates achieving full marks did use Oldham and Brass and used aspects of it as an example, such as the type of workers in their study (newspaper) and where the study was located (United States).

Section B

Question 9

(a) Many candidates decided to conduct an experiment and wasted time writing about IV, DV and other aspects of experiments which were not required because the question asked for a **questionnaire**. A second flaw in many answers is that they did not use *open questions*, again a requirement stated in the question. Yet another flaw was that candidates often assumed that the cause of anxiety was examinations, presumably because the participants were university candidates. However, generalised anxiety is when the cause of the anxiety, perhaps panic attack, is not known. A number of candidates applied the GAD-7, which uses *closed* questions, and this was creditworthy if it was used alongside open questions. The GAD-7 alone could only receive minimal credit. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

© 2024

(b) (i) For psychological knowledge candidates often referred to diagnostic criteria from ICD-11 for generalised anxiety disorder. Many candidates stated that 'it is where the cause of anxiety is unknown' despite focusing on 'exam anxiety' in their **part (a)**. Candidates also wrote about the GAD-7 which is specifically for testing generalised anxiety and received credit.

(ii) Candidates achieving full marks for this question often outlined a characteristic of generalised anxiety which they had included in **part (b)(i)** and which they had included as one of their open questions in **part (a)**, thus successfully linking their plan **part (a)** and **part (b)(i)**. Another creditworthy feature was to use the GAD-7 questions and convert them into open questions, or to use a GAD-7 closed question with an additional open question.

(c) (i) Most candidates answered this question part incorrectly. The question asked for open questions, so answers to this question part should have focused on how to score/interpret open questions. Candidates choosing to focus on interpreting the closed questions of the GAD-7 achieved no marks because how they are scored is known knowledge and not a candidate applying knowledge to their plan. Candidates were awarded marks for explaining that two (or more) judges would be needed to achieve agreement over the answers to the open questions.

(ii) For this question part the candidates explaining how they interpreted responses to open questions (in **(c)(i)**) correctly stated that there might be ambiguity or bias in the judging and that an inter-rater check of consistency would be needed, assessed with a correlation.

(iii) The term 'questionnaire technique' still confuses many candidates. It is not the use of closed and/or open questions, which is question format, rather it is whether the questionnaires are pencil/paper (face-to-face) or postal or online. Candidates answering the question correctly gave an appropriate reason but often failed to relate the reason to their plan.

Question 10

(a) There were three types of answer in response to this question: **(i)** those who knew what an explorer type of shopper was and investigated this using postal questionnaires. Such answers often received high marks because they addressed the question; **(ii)** answers which investigated the explorer type of shopper but by interviewing them in the supermarket rather than using a postal questionnaire; **(iii)** answers which did not know what an explorer type of shopper was and who did not use a postal questionnaire. Candidates should have studied two options and so if the explorer type is not known then the question from the second option should be considered. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) For psychological knowledge the most appropriate research would be that by Gil et al. (2007). This is an 'e.g.' study and an alternative study could be substituted. However, movement patterns (four types) and behaviour patterns (five types including explorer) are listed on the syllabus and so should be studied. Answers that scored highly in this question part were those who could describe the behaviour patterns in detail, those who could not describe any behaviour pattern at all tended to receive no marks.

(ii) This question part required an explanation to show how what was described in 10**(b)(i)** informed the plan in **part (a)**. For example, in **part (b)(i)** the explorer behaviour pattern should be described. In **part (b)(ii)** it should be explained how the explorer type informed questions that were included on the postal questionnaire.

(c) (i) Many candidates chose an opportunity sampling technique and the reason for this choice was that they were available in a supermarket. This was correct to a point but fell down when candidates went on to describe that they would be interviewed, forgetting that the required method was a postal questionnaire. An opportunity sample might not be the best technique to use because participants did not need to be in a supermarket at all because they are doing a postal questionnaire.

© 2024

(ii) For candidates choosing opportunity sample in **(c)**i, they struggled to find a suitable weakness, often nothing more than 'there might be researcher bias'. If a different sampling technique had been chosen for the postal questionnaire then there would have been more points to consider for weaknesses.

(iii) Many candidates chose to gather quantitative data and often gave an appropriate reason for that choice. However, most candidates could not be awarded full marks because they failed to relate their reason to the quantitative data they planned to gather in their study.

Question 11

(a) Some candidates answered the question and wrote excellent answers showing good planning and application of appropriate methodology. However, there were many candidates who could not be awarded marks for two main reasons: **(i)** they ignored the question and did not use a *telephone interview* to gather data; **(ii)** they ignored the question to interview the children *when they are adults*, instead planning a study involving a sample of children and in effect describing what was done in the Tapper et al. study. A few candidates then added a comment at the end 'and then interview them as adults'. Candidates must read the question carefully and plan a study based on that instruction. Any deviation from it can only be awarded bottom band marks. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) For psychological knowledge the most appropriate research was that by Tapper et al. (2003) using positive reinforcement and the 'Food Dudes'. Some candidates described this in detail and were awarded full marks. Others described some aspects of the study and others could not write anything at all.

(ii) Candidates being awarded full marks wrote about how they used their knowledge of the Tapper at al. study to inform their plan. At a simple level, this could involve asking the adults the closed question 'Do you still eat more fruit and vegetables? yes/no' in the telephone interview. Rather than answer the question, some candidates explained their methodology, such as their sampling technique.

(c) (i) 'Question format' is whether the questions are open or closed and because the method is a telephone interview the logical choice would be closed so the interviewer could easily record the answers to yes/no questions such as 'Do you still eat more fruit and vegetables? yes/no'. An answer like this would be awarded 2 marks because there is a reason for the choice of closed question and it is related to the study.

(ii) Following from **(c)(i)** a logical weakness of asking closed questions would be that 'the reason why they eat more or less fruit and vegetables would not be known (was it due to the intervention or not)'. This answer would be awarded 2 marks because there is a weakness (reason why not known) which is related to the study (fruit and vegetables).

(iii) Some candidates do not know what an interview format is, often explaining that it was a telephone interview, which is the interview technique. Candidates awarded full marks explained that to gather data to allow comparison a structured interview was used so all participants would be asked the same questions in the same order. Those linking this to their plan were awarded full marks, those not linking to their plan were not.

Question 12

(a) Most candidates constructed a plan based on slow and rapid rotation, but many did not know more than the terms themselves, so their plan was often confused. The question stated 'negative effect on the health of workers' and some candidates did not do this, instead having their DV as the number of accidents, or the lack of sleep. Better answers planned a field experiment with an IV of slow and rapid rotation and DV as the number of consultations with a doctor about more serious illnesses. Controls were often the use of the same workers in the same industry. A number of candidates replicated the Gold et al. study, using nurses, their rotating shifts and accidents and sleep. Candidates should not describe a study they have learned; the requirement is to plan their

© 2024

own, individual study based on psychological knowledge. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) Relevant psychological knowledge here was shiftwork and some candidates described the different shift patterns, such as the 'metropolitan rota' in detail. Some candidates described the study by Knutsson (2003) who reviewed the effect of shiftwork on health and others described the Gold et al. study on nurses. Marks were awarded according to the quality detail and accuracy of their answers and, for 4 marks, this question part required more than two sentences.

(ii) Many answers were awarded full marks for linking their **part (b)(i)** answer to their **part (a)** plan. Other candidates did not answer the question at all, sometimes describing their **part (b)(i)** in more detail, sometimes describing some aspect of their plan and sometimes the question was not answered at all.

(c) (i) This question asked candidates about the controls they had applied to their plan. Although applying controls is fundamental to any experiment, there were many candidates who did not do this. Before starting to write candidates are advised to read the question, *including the two bullet points*, which are essential to include in their plan because questions in this **part (c)** sub-section will focus on them. If a candidate arrives at question like this on controls, there is no reason why they cannot add to their **part (a)** answer. Adding a control only in this question part, but which is not in the plan will not attract marks.

(ii) Candidates often applied generic weaknesses of controls, such as demand characteristics being more likely as knowledge of being in a study becomes more evident, but rarely was this applied to the controls in the **part (a)** plan, if there were controls there at all. It is recommended that candidates read all question parts before starting their answer so they can plan ahead and ensure their answers to all question parts are coherent. In other words, look at this question part and include a control in the plan *that has a weakness*.

(iii) The choice of dependent variable should be one that tests the question, in this instance, the negative effect of shiftwork on health. Following on from AS knowledge the DV should be fully operationalised in order for full marks to be awarded. Answers often were as brief as 'the DV is health' which was too vague for marks to be awarded.

4Uadmission

© 2024

PSYCHOLOGY

Paper 9990/43

Specialist Options: Application and Research Methods 43

Key messages

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

General comments

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

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

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

Section A

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

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

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

Section B

Candidates should only answer one question from this Section.

4Uadmission

© 2024

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

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

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

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

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

Answers to **part (a)** questions in this section should include an appropriate plan, have applied a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Many answers listed features such as 'I would have a random sample' and 'It would be an independent measures design' without explanation of why it would be a random sample, or how it would be obtained. Elaboration of these general sentences should be included.

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

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

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

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

Comments specific to questions

Section A

Question 1

(a) (i) One mark was awarded for candidates stating that a clinical interview is an interaction between a medical practitioner and a person/patient. Additional marks were awarded for stating that this would take place in a medical setting, where verbal and non-verbal exchange took place and the purpose of which was to diagnose, assess and treat. Some candidates incorrectly assumed that a clinical interview was part of a 'study' with a researcher asking questions.

© 2024

(ii) Participants were interviewed by a second psychiatrist (as the question stated) mainly to confirm the data gathered in the first interview. Many candidates understood this and were awarded full marks. However, some candidates wrote as if this was the first interview to gather initial data and so could not be awarded full marks.

(b) Nearly all candidates wrote about researcher bias, or lack of consistency between the two psychiatrists and were awarded one mark. Some candidates correctly applied their point to the key study and were awarded a further mark, but some candidates did not do this and could only be awarded one mark.

(c) Most candidates had little difficulty in providing a strength and a weakness but frequently failed to locate the strength or weakness in the context of the study. For example, a commonly stated weakness was 'participants might not want to answer because of social desirability'. This is far too general, and it needs an explanation to be awarded marks. Writing in addition, for example, 'because they are embarrassed to acknowledge their family history of depression' makes a good answer worth two marks.

Question 2

(a) Some answers were vague with no more than 'the urge to set fires' but were awarded one mark. Candidates adding more detail, such as 'persistent fascination or preoccupation', 'experience of pleasure, excitement, relief or gratification during or after fire setting' were awarded full marks.

(b) Some candidates suggested giving pyromaniacs questionnaires or interviewing them, but as both these are self-reports no marks were awarded (the question stated …*other than by self-report)*. Answers scoring full marks suggested conducting a covert observation (for example by a family member) of a pyromaniac with a checklist of typical behaviours and recording the frequency of different fire-related activities.

(c) There were some excellent strengths included such as a self-report giving the tester a measure of the extent of the problem or that a self-report measure can help to understand the characteristics of pyromania such as its addictive nature. However, whilst some candidates related these points to pyromania and were awarded full marks, others did not and so marks were restricted. Full marks are awarded for a strength (one mark) plus applying it to the sub-topic (+one mark).

Question 3

(a) (i) Any two features of the sample were awarded marks, which included: 180 customers, 118 female 62 male, aged 16 – 80 years, lived in/around Lund, Sweden. Most candidates were awarded two marks. Some candidates wrote incorrect numbers, some described the supermarket (not a feature of the participants), and some believed the study was conducted in Germany, probably confusing this with the Becker et al. key study.

(ii) Correct identification of the sampling technique, opportunity sampling, achieved one mark and elaboration of this, such as 'approached as they passed the tasting stall' or 'passed the stand in the supermarket' was awarded an additional mark. Many candidates did not know the answer because they guessed at 'volunteer sample' or 'random sample'.

(b) Correct suggestions could only be 'volunteer sample' or 'random sample', which most candidates did and by stating how these techniques could gather participants, were awarded full marks. Some candidates made a correct suggestion and with no elaboration were awarded one mark. Some candidates incorrectly suggested an opportunity sample and could not be awarded any marks.

(c) Like other **part (c)** questions, candidates gave a strength and weakness without relating it to the question, and sometimes answers were very vague. For example, 'there is no selection bias by researchers'. Answers like this receive no marks because there is not even an indication as to what sampling method is being commented on and Examiners cannot make assumptions for candidates. An strength awarded full marks might be '…because the participants are chosen randomly, there is no selection bias by the researchers in their study of tasting jam and tea in a supermarket.'

Question 4

© 2024

(a) A few candidates gave tautological answers which simply re-wrote the question such as, 'a customer-focused sales technique focuses on the customer' and these received no credit. Most candidates provided very good answers which showed good understanding of the concept. Such answers mentioned considering the needs of the customer, letting the customer talk and gathering information about customer preferences and the uniqueness of each customer.

(b) One effect of a customer-focused' sales technique is that the customer feels valued, that they were listened to, and that they are happy. This means that they are likely to return to buy another new car and probably recommend the dealer to their friends. Answers like this were provided by most candidates who were awarded full marks.

(c) Many answers included a strength and a weakness of competitor-focused techniques. A number of candidates linked their points to buying a car, following on from **(b)**, and others chose different examples. Overall understanding of these techniques was very good, perhaps because candidates had direct knowledge of how sales techniques are applied.

Question 5

(a) (i) This question asked *how* participants were recruited, not to identify the sampling technique. As the study states: 'a random number generator was used…' and any elaboration of this, such as 'four patients from each surgery', 'from an inner London general practice', or 'over a four-month period' were awarded the additional mark because they showed knowledge of the key study. Candidates stating nothing more than 'it was a random sample', or any other sampling technique scored no marks.

(ii) This question part was answered correctly by nearly all candidates who knew that 'face-down cards were turned over by the doctor as the patient entered the room' to determine whether the directing or sharing style was used.

(b) Many candidates were able to suggest that not randomly allocating patients to conditions could lead to bias and so questioning the validity the study. Often included was the comment that doctors might select a style for a patient because they knew the patient, or the patient 'looked' like a 'sharing' sort of person.

(c) Some candidates showed good insight by realising that random allocation prevents both participant bias and doctor (in this instance) bias. This, when linked to practitioner style and patient satisfaction could achieve full marks.

Question 6

(a) Attention diversion to manage pain was clearly explained by most candidates who often suggested that it could be passive or active (such as doing a puzzle). A few candidates confused it with non-pain imagery or cognitive redefinition.

(b) There are two alternative (as defined by the syllabus) treatments for pain which are acupuncture and stimulation therapy (such as TENS). Some candidates only identified one of these for 1 mark, but others provided an explanation and/or example for 2 marks. A few candidates suggested taking medicine/drugs but this is not a syllabus defined alternative and could not be awarded marks.

(c) There were some excellent answers with strengths focusing on the ease of use for patients and weaknesses focusing on the difficulty of concentrating enough to divert attention, or that the technique applied more to mild pain and not to chronic pain.

Question 7

(a) (i) In order to be awarded marks, answers to this question needed to show some knowledge of the Swat key study. Swat defined accident frequency as 'the number of *injury* accidents with sick leave days off work per 100 employees in a year. Candidates rewording the question and writing 'accident frequency is the number of accidents' scored no marks.

(ii) To be awarded marks for this question candidates needed to identify the two industrial plants in the Swat study which had the highest accident frequency rate. The two were 'foundry' (5.9) and 'meat

© 2024

processing' (2.8). Any other industrial plant, such as 'machinery' (2.1), or 'furniture' (2.3) could not be awarded marks.

(b) Many candidates correctly wrote about a psychological technique specifically to reduce accidents, such as using a token economy system as outlined by Fox et al (1987). Others applied their more general psychological knowledge and wrote about positive reinforcement if correct safety procedures are followed. A third type of answer often scored no marks with comments such as 'tell them off if they make a mistake'.

(c) The most common strength was that 'it is quantitative and so can be compared' (which is correct for one mark) but if there was no elaboration beyond this and no mention at all about accidents then no further marks could be awarded. The common weakness about accident frequency was that many accidents go unreported or that only major accidents are recorded, not minor ones.

Question 8

(a) Nearly all candidates could outline Maslow's 'physiological needs' and 'social needs' with little ambiguity. Some candidates described the other needs and even the hierarchy even though no additional marks were available for doing this.

(b) Candidates frequently suggested 'doing a questionnaire' for which no marks could be awarded because this is far too vague. The question technique (face-to-face, postal, online) and/or the questionnaire format (open and/or closed questions) should always be included. Candidates being awarded full marks suggested doing a questionnaire with open questions, which would allow the person to talk about their achievement of social needs with colleagues in the workplace.

(c) Only a few candidates demonstrated good knowledge of telephone interviews because most answers were anecdotal, and so their marks were restricted. One strength is that a potentially diverse sample from anywhere in the world could be interviewed, and a weakness is that there is no non-verbal communication, which is often a useful indicator in an interview. Very few candidates related this to need theories of motivation.

Section B

Question 9

(a) Many candidates planned an experiment and the IV was usually MAOIs compared to SSRIs. Some candidates vaguely had 'antidepressants and controls' being unable to name the anti-depressants. The design was commonly an independent design, but some candidates incorrectly claimed repeated measures because each participant repeated taking the same drug. Some candidates had a long list of appropriate controls, such as taking no other drug during the trial and taking the drug for 6 months. Some candidates included a control group with no medication, but this would not allow determination of which antidepressant was more effective. DVs were often not operationalised although some good answers used the Beck Depression Inventory (BDI) to measure the level of depression before and after the trial. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) For psychological knowledge many candidates described the way in which both SSRIs and MOAIs work. Some described the characteristics of depression, and some described the features of the BDI.

(ii) A number of candidates linked what they had written in **(b)(i)** with what they had planned in **(a)**. For example, they might describe SSRIs and link that to why in their plan participants would take one tablet daily or why they conducted the study over a six month period as they needed to allow time for the medication to work. However, some candidates incorrectly described an aspect of methodology (with no link to **(a)**), some continued what they were describing in **(b)(i)** and some candidates could not provide an answer.

(c) (i) Most candidates answered this question part correctly by giving a reason why they used an independent design. Those using repeated measures in their plan struggled to give a reason for its

© 2024

use. A few candidates incorrectly thought 'experimental design' meant 'type of experiment' and explained why they had conducted a laboratory or field experiment.

(ii) The most common response was that an independent design requires more participants. This is not necessarily the case; the design does not determine the number of participants: if the sample size is ten participants in each group, then ten it is. Further, there is the counter argument that more participants might be needed but for less time; a repeated design might have fewer participants, but for double the time. Those using the 'more participants' argument often failed to relate it to their plan.

(iii) The choice of type of experiment produced interesting responses. Some planned a field experiment because participants would be taking the drug in their own home; others planned a laboratory experiment because participants would go to the laboratory once a day to receive their medication. Either argument was legitimate if a reason was provided and full marks could be awarded if it was related to the plan.

Question 10

(a) This question required use of a questionnaire with closed questions. Most answers were good but often did not explain in enough detail for full marks to be awarded. For example, writing 'I chose an opportunity sample' without a comment on how this would be done. Similarly, 'the questionnaire has a four-point scale' without saying what the four points are. Details of the questionnaire technique were commonly absent with nothing more than 'I then gave participants the questionnaire' rather than explaining why they had chosen a particular technique. A simple 'I gave participants a questionnaire and a pencil so they could fill it out before they left the supermarket' would be sufficient. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) For psychological knowledge most candidates could explain multiple unit selling although sometimes examples given in support were ambiguous. The study by Wansink et al. (1998) was also described with some candidates quoting the Wansink examples and data (such as 45 per cent more sales when using multiple unit pricing).

(ii) This question part required an explanation to show how what was described in 10**(b)(i)** informed the plan in **part (a)**. The best format is: **part (a)** should include a specific example of multiple unit pricing that can be used on participants. **Part (b)(i)** should describe what multiple unit pricing is and an example of it (perhaps from Wansink et al). Question **part (b)(ii)** then explains the reasoning behind the example used in **(a)** because it is based on the knowledge described in **(b)(i)**.

(c) (i) Question technique concerns how the questionnaires are administered to participants which could be online, postal or pencil/paper (face-to-face); questionnaire format is about whether questions are open and/or closed. Candidates opting to give a reason for the latter were awarded no marks. Any of the three techniques were appropriate and to answer questions participants could do the questionnaire on-line at home; they did not need to be in a supermarket before answering the questions.

(ii) Many candidates stated nothing more than 'the reason why is not known' which is far too vague because this could apply to anything. For marks to be awarded there must be some comment about the study that has been planned. For example, 'the reason why people opt to choose multiple units (e.g. 3 for $30) rather than a single one (1 for $10) will not be known.

(iii) Many candidates could identify an appropriate descriptive statistic. What they could often not do was to give a reason why they chose that particular statistic. For example, a candidate might write 'I chose the mean to analyse the data' without giving the reason why the mean was chosen over the median or mode. Answers often failed to be related to the plan.

Question 11

(a) Many candidates decided to use every aspect of the McKinstry and Wang study and often their answers were merely a description of that study. Some candidates even included the results found

© 2024

by McKinstry and Wang. Candidates should plan their own investigation that may use one or two aspects of a published study. For example, photographs could be used but the photographs should not be exactly the same as the original. Also, real people could be used rather than photographs. Candidates must show that they can think for themselves and plan their own individual investigation. Some candidates did plan something original, and these answers were interesting to read. Another weakness in answers was that candidates decided not to use a face-to-face interview, instead using a completely different method. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) For psychological knowledge the most appropriate research was that by McKinstry and Wang (1991) who compared five styles of a doctor's dress, including a white coat. Notably that is an 'e.g.' study meaning that an alternative could be used. However, candidates either described the McKinstry and Wang study or no study at all. No candidate described an alternative study.

(ii) As applies to all **(b)(ii)** questions candidates should explain how they used the knowledge described in **(b)(i)** to inform their plan. Some candidates did this very well and were often awarded high marks. Some candidates did not answer the question correctly and could not be awarded marks.

(c) (i) Candidates are advised to read all parts of the question before starting their **part (a)** answer, paying particular attention to the bullet points, because those points will be asked about in **part (c)** questions. In this instance candidates got to this question realising that they had not addressed how the data their plan gathered would be interpreted or scored. This meant that answers were often vague and not linked to the plan.

(ii) Knowledge about face-to-face interviews is very general given the answers to this question. This method should be given just as much attention as any other and so candidates should know at least two strengths and weaknesses. Comparisons with other types of interviews such as telephone should also be known. Most candidates simply stated that 'participants might lie to the interviewer' which might be the case, but they are perhaps more likely to tell the truth because they are face-to-face.

(iii) Many candidates are still confused about the difference between interview *technique* and interview *format*. In this instance the question asked about interview format which refers to structured, semi-structured, or unstructured interviews. Again, candidates should have paid attention to the 'interview format' bullet point and should have used a specific format (structured, semi-structured, or unstructured) in their plan, rather than, for example 'I would ask five questions'.

Question 12

(a) Most candidates planned an experiment, and many applied a number of features specific to experiments, such as IV, DV, controls, experimental design and type of experiment. An interesting range of physical work conditions was included, such as seat design in office workers or the colour or type of lighting (LED versus halogen), although some candidates still used temperature as an IV despite the question telling them not to. Candidates must remember to include a range (four or five) of specific (to the named method) methodological features, each of which should be explained fully to show good understanding. Candidates should also include appropriate 'general' methodological features such as ethics, sample, sampling technique and location of the study. Candidates must remember to include the two bullet pointed features in their plan.

(b) (i) The syllabus states 'impact of physical working conditions on productivity' and so candidates should have relevant psychological knowledge of several conditions, most logically lighting, temperature and noise depending on the industry, but also office design (e.g. open plan) or other features could be included. The study at the Hawthorne plant could also be included, best described by Kompier (2006).

(ii) There were two types of answer: candidates who explained and related what they had planned in **(a)** to what they had described in **(b)(i)** and so could be awarded full marks and candidates who did not answer the question at all, or those who provided a partial answer.

© 2024

(c) (i) This asked about the choice of DV and some candidates merely referred to 'productivity'. Any DV must be fully operationalised which was a 'free choice', given there was no specific industry identified in the question. A DV could have been 'the number of cuddly toys produced each hour'.

(ii) The choice of type of experiment was often a field experiment, conducted in some factory, which is logical given the nature of the question. Less logical were candidates who planned a laboratory experiment which is a more controlled, but a less realistic situation. Importantly, marks were awarded for the reason given, whichever type is used, and if the reason is appropriate and linked to the plan then full marks can be achieved.

(iii) One or more ethical guidelines must be *included in every plan whatever the question* and so giving a reason for one of these guidelines should have been straightforward. However, commonly there was nothing more than 'I debriefed participants', which is not a reason *why* this is done, and such a brief comment could relate to anything rather than the specific plan of the study.

4Uadmission
SINCE 2020

© 2024