PSYCHOLOGY

Paper 9990/12

Paper 1 Approaches, issues and Debates

Key messages

Candidates need to know all components of the study as listed in the syllabus. Questions can be asked about any part of a 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 or a named issue to be included. 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 last one of these about the named issue. Credit is limited if the named issue is omitted. In addition, if the candidate is required to explain a similarity or a difference then they must explicitly do so.

Candidates need to be careful about how they present the results of studies. For example, they need to know if the results are about how many participants performed a task correctly or on how many trials the participant was correct. This can have a large impact on the interpretation of results and whether a response can gain credit or not.

Candidates also need to engage with any stimulus material presented in a question (e.g. 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.

Candidates need to be able to know about real-world applications for all core studies. To show understanding, answers need to tell the Examiner what the application is, based on the particular core study, and then how the study will be useful.

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

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

Questions about the 'psychology being investigated' require answers that are generically describing concepts and theories that are core to the study. Minimal credit can only be awarded for answers which describe the processes of the study itself rather than a generic description of a theory or concept.

There is enough time for answers to be planned to ensure that the response given by a candidate is focused on the demands of each question.

General comments

The marks achieved by the candidates sitting this examination covered a wide spread of possible marks. Some candidates provided a range of excellent answers to many of the questions and could explain psychological terminology well, providing evidence that they were prepared for the examination. There is no evidence that candidates had not learned the new studies that form the 9990 syllabus. This was also evidenced by very few blank answers.

bestexamhelp.com

© 2020

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.

Comments on specific questions

Question 1

(a) Many responses to this question were correct (independent groups). Some responses named the type of experiment or sampling technique used which was not the question set. Some responses included a reason why independent measures may have been chosen but the question did not require this.

(b) There were many good responses to this question with clear outlines of materials that were provided to participants in the doodling condition. Common materials included the paper with circles and squares on it plus the pencil provided.

(c) There were many incorrect responses to this question. The majority of responses gave the result for false alarms for names whereas the question was asking about the false alarm for places.

Question 2

(a) Many responses could clearly identify a psychological concept that was being investigated in the study by Saavedra and Silverman. Common choices were included phobias, evaluative learning and operant conditioning. Whilst some responses could then outline these in general terms, many responses then instead focussed on description of the study.

(b) Reponses to this question were varied. Some responses could outline one result from the posttreatment assessment session, including him no longer fulfilling the criteria for a specific phobia. Many responses outlined a result from the therapy sessions which could not gain credit as the question was about the posttreatment sessions.

Question 3

(a) Some responses could outline one result from Alex using familiar objects, and present data. Some responses gave a conclusion about what Alex was capable of which could not gain credit.

(b) Stronger responses could identify a methodological strength of the study. Popular choices were about standardisation and the use of controls. Some candidates mixed up these two concepts and could only be awarded partial credit (standardisation aids reliability; controls aid validity). Some responses were able to fulfil the ‘of this study’ part of the question by providing an example from the study to explain why it was a strength.

(c) Popular choices included housing, socialisation and caging with some well thought through answers. Responses about boredom were awarded partial credit as examples of Alex being bored are not presented in the Pepperberg paper. Responses needed to be about aspects of ethics that we know happened. Responses about not being the wild/natural habitat are not relevant to the study in terms of ethics.

Question 4

(a) Stronger responses to this question could clearly describe the learning task used. However, there were a proportion of responses that described the shock instructions and not the learning task. It is important that candidates are familiar with the terminology used in the original articles as question can be based around these.

(b) Many responses could identify that the draw had been rigged so that the participant was always the teacher.

© 2020

Question 5

The majority of responses to this question could suggest at least one instruction that Sai could have used to replicate the part of the study by Bandura presented in the question. However, there were a significant amount of responses that did not give instructions for what was the first part of the procedure; the experimental room. There were many responses outlining the aggression arousal part of the study (second part) and the test for delayed imitation part of the study (third part). This question is good example where candidates need to read carefully ensuring that their responses are focused on the part of the procedure asked in the question.

Question 6

(a) Reponses to this question were mainly strong. Many responses could identify two characteristics of Group 1. Characteristics refer to the participant variables that the group has. Some responses described features (e.g. sampling technique) which could not gain credit for this question.

(b) There was evidence of strong responses here with candidates able to clearly describe how Group 1 was recruited in the study by Baron-Cohen et al. However, there were responses that described the recruitment of a different group or a different study entirely.

(c) Responses clearly discriminated between candidates who knew what a conclusion was compared to those who did not. A conclusion is a generic commentary summarising what was found in the study *based on the results*. Many responses gave results from the study. These could not gain credit as the question required a conclusion. Popular conclusions focused around Theory of Mind, improved validity of the Eyes Test and gender differences in levels of autism.

Question 7

(a) The large majority of responses could clearly outline one aim of the study by Laney et al. Some responses focused on subsequent eating behaviour which could not gain full credit as this was never investigated.

(b) There were many examples of real-world applications based on the findings of Laney et al. Popular choices included helping fussy eaters, changing diets for people with obesity and aiding the diet of patients being treated for cancer. Many responses could outline what the application would be and how it would be implemented to gain both available marks.

Question 8

(a) Most responses could outline at least one assumption of the social approach. The main choices were the effects of groups on behaviour and the effects of individuals on behaviour. Some responses gave clear examples for each assumption. These examples were a mixture of ones from core studies or ones created by the candidate themselves. Both approaches were equally valid.

(b) Stronger responses could clearly choose relevant results and use them to explain why the result supported the situation debate and did not support diffusion of responsibility. A popular choice for the situational debate was about the differences in help based on the state of the victim. Many responses could then explain how the result was based on situational factors. A popular choice for the diffusion of responsibility part of the question was about group size and speed of helping. Many responses could then explain how that result did not support the concept. However, there were some responses that claimed the model received help and these could only gain partial credit as it was the victim that received help.

Question 10

The strongest responses evaluated the study by Canli et al. in depth and in terms of two strengths and two weaknesses with at least one of these points covering the named issue of laboratory-based research. Common choices included types of data collected, reliability, generalisability and ethics. These strong responses could explain why an element of the study was a strength or a weakness using specific examples from the study by Canli et al. explicitly to support their point. These answers tended to score Level 4 marks. Candidates need to ensure that they follow the demands of the question, covering two strengths and two weaknesses all in equal depth. Some responses did cover the four evaluation points but were brief or did not use the study by Canli et al. as examples which meant the response scored in the lower bands. Other

© 2020

responses included three evaluation points that were thorough, logical and well argued with a fourth point that was brief which meant the response did not reach the top band in the main. Candidates need to know that any description of the study does not gain credit in these type of questions as it is testing their evaluation skills only. In addition, some candidates were following a GRAVE approach to this question (Generalisability, Reliability, Application, Validity, Ethics). Therefore, some candidates were producing prepared essays for Canli et al. without one of their points being about laboratory-based research. 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 what has no context cannot get above a Level 1 mark. To improve on this question, candidates need to plan carefully, choosing two strengths and two weaknesses with one of these being the named issue. Each strength and weakness should be of equal length with an explanation as to why it is a strength or weakness with example(s) from the study to show clear understanding. These are the requirements of a Level 4 response.

© 2020

PSYCHOLOGY

Paper 9990/22
Paper 2 Research Methods

Key messages

- This is a question paper about research methods, which requires candidates to answer a range of question types, including ones about the core studies in relation to research methods, terms and concepts used to describe or evaluate research methodology, and application of this knowledge to both familiar and unfamiliar contexts. Candidates demonstrated good understanding in many of these areas although some common flaws were evident in each of these skills. It is therefore essential that candidates are prepared for the skills of recalling concepts and of using this knowledge.
- Practising the application of ideas, especially to novel scenarios and in learners' own practical activities, is important to success on this paper:
  - Candidates need to be able to apply research methods terms and concepts to scenarios presented in questions. These can include, for example, planning, criticising or developing designs or analysing data
  - Candidates must take note of questions which indicate the need for a link. When a question says 'in this study', or makes direct reference to the scenario, responses must go beyond simply describing or evaluating, they must contextualise the answer ***in a relevant way***. Candidates therefore need to be prepared for questions using this format and practice can help them to learn both how to extract relevant ideas and how to make novel suggestions based on scenarios.
- **Question 10** in this paper requires candidates to produce an original design for a novel research question; this 'creative' process requires practice. Furthermore, to learn to identify flaws in a design (whether their own, as in **Question 10**, or one from a novel scenario for example in **Section B**) also relies on having had experience of practical problems in conducting studies. This is a high-level skill, and can be developed through practical work with designing and conducting small studies in class or through the discussion of novel scenarios. The overall format of **Question 10(a)**, and the nature of the mark scheme, is consistent between papers and years. Therefore, it is helpful to prepare candidates with an overall structure, which can be closely tailored to the requirements of an individual question, such as the required research method and the scenario.

General comments

In general, candidates were able to access marks across the whole paper. However, very few were consistently able to access the marks for linking their response to the scenarios, which limited performance overall. Some candidates demonstrated a good grasp of a range of psychological concepts and were able to access the basic marks with these.

Candidates across the ability range were able to demonstrate some knowledge of a range of aspects of research methods in this paper. Success was greater on questions such as **Questions 1(a)**, **3**, **4(a)**, **4(b)** and **7(b)(i)**, **(ii)** and **(iii)**, than on **Questions 7(c)** and **10**. This examination tested a cross-section of psychological skills and some candidates showed limited success, such as in **Question 3**, where some responses made incomplete statements about lack of generalisability and **4(a)(ii)** where candidates were unable to follow through the effect of individual difference between participants to the effect of this on group differences in an independent measures design. Note also that access to full marks on **Questions 7(a)(ii)** and **8(c)** and **(d)** was limited by omitting to respond to the instruction to link to the scenario. Finally, there was considerable confusion between privacy and confidentiality in **Question 9**, which primarily resulted in low scores in **Question 9(b)(i)**.
Question 10 was sometimes well answered although many responses were experimental rather than correlational and lacked linear measures of the two covariables.

© 2020

Comments on specific questions

Section A

Question 1

(a) This question part was generally well answered although there was a significant minority of responses that earned no marks because they suggested that data was collected by 'field experiment'.

(b) Most responses to this question were generic, that is they did not include a link to the study by Piliavin et al., as required by the question. Such partial answers could only earn limited credit.

A small number of responses described quantitative data as 'accurate'. This is too imprecise to be creditable.

Some responses attempted to suggest that an advantage was that it was 'numerical'. Since this is descriptive it is not answering the question.

Question 2

There were some answers that clearly demonstrated a good level of understanding of the concept in alternative ways, such as 'We can generalise because the participants were all right-handed, and the majority of the population is right-handed.'

There were a number of candidates who attempted to define generalisability with the word 'generalisable'. Such circular responses were not answering the question, so could not earn credit. Another common way in which candidates provided incomplete responses was to only partially explain examples, such as identifying that the sample was all 'female and right-handed' but then omitting to state that findings from this group were therefore not generalisable to males / left-handed people.

Question 3

Some responses did not indicate the *effect* on the DV. Such answers tended to be repetitious.

Question 4

(a) (i) This question was very well answered. Candidates demonstrated a good understanding of the design and were able to name it accurately.

(ii) Many responses to this question were too simplistic, not going beyond a statement of the 'disadvantage is individual differences'. Few candidates understood the disadvantage well enough to explain that the reason why such differences matter is that they can create (artificial) differences between *groups*, i.e. levels of the independent variable.

(b) Although there were many good answers to this question part, one common error, leading to partial marks, was to repeat the stem, e.g. to say they were 'deceived about the false memory'. This is not fully answering the question as it only identifies an ethical problem (i.e. deception). Such candidates need to go on to describe the nature of the deception used in the creation of the false memories.

Question 5

This question was not well answered. Candidates often began correctly, for example identifying that this controlled for participant variables, but then incorrectly continued to say, for example, that all the aggressive children would be in one group (e.g. to encounter the non-aggressive model) and all the non-aggressive children in another group.

Additionally, some candidates merely elaborated on the question, saying that this was to allocate them to groups without specifying how.

© 2020

Question 6

This question was typically well answered, both in terms of the different ways suggested to measure dependent variables and the level of examples given. Responses were typically accurate and relevant. One common error arose if candidates chose quantitative and qualitative as their 'ways'. They tended to score well in their descriptions relating to quantitative data, but in their response linked to qualitative data, it was more difficult to give genuine examples of measurement of a DV.

Section B

Question 7

(a) (i) A significant minority of candidates answered this question part incorrectly, often repeating a technique from the stem.

(ii) This question part was moderately well answered although many responses lacked a link to the study that was required for full marks.

(b) (i) There were a significant number of incorrect responses to this question part. This was most often because candidates offered questions that could be answered with a single word, i.e. which were in fact closed questions. An open question must be capable of eliciting a range of individual responses.

(ii) There were a significant number of incorrect responses to this question part. This was most often because candidates attempted to write closed questions but did not provide answer options.

(iii) This question part was well answered.

(c) This question part was fairly well answered, with most candidates gaining at least partial marks. The most common problem with the responses was that they tended to be generic rather than linked to Jenny's study of Sven.

Question 8

(a), (b) These question parts were very well answered.

(c) This question part was fairly well answered, with most candidates gaining at least partial marks. The most common problem with the responses was that the advantage tended to be generic rather than linked to investigating healthy eating. In other words, the advantage was applicable to healthy eating but this had not been explained.

(d) This question part was fairly well answered, with most candidates gaining at least partial marks. The most common problem with the responses was that the disadvantage tended to be generic rather than linked to investigating healthy eating. In other words, the disadvantage was applicable to healthy eating but this had not been explained. For example, 'they may lie' is a disadvantage. 'They may lie about healthy eating' does not explain this disadvantage. In contrast 'They may lie by saying they ate salad when they had a burger' explains what the disadvantage is in research on healthy eating.

Question 9

(a) This question part was very well answered, with many candidates gaining at full marks. Two common problems were that the axis labels but not headings were present, and that the bars of the bar chart were not separated.

(b) (i) Many responses to this question part demonstrated that the concept of privacy is not well understood. This was often mixed up with consent or with confidentiality.

(ii) Responses to this question part were better than those for **part 9(b)(i)**, with many candidates gaining at least partial marks.

© 2020

(c) This question part was well answered by many candidates. However, a minority of responses indicated good understanding of the meaning of a structured observation, but then lacked a link to Silas's observation. In such cases the response was limited to partial marks.

Section C

Question 10

(a) Many candidates designed experimental studies (mainly laboratory or field experiments with an independent measures design) and incorrectly described levels of the independent variable, how participants were allocated to these and how their procedures differed. This was not creditworthy and cost these candidates valuable time. Although their measurement of their dependent variable could earn marks, this was often simply a categorical measure of 'did or did not doodle', producing data that could not be correlated. The suggested measures of understanding described were generally relevant, however.

(b) Many responses to this question part scored zero, for one of several reasons. Candidates who had described an experiment in **part (a)** typically offered irrelevant criticisms. Candidates who had correctly described a correlation, however, did not necessarily offer changes to their own procedure, suggesting instead entirely different studies.

For candidates trying to answer the question appropriately, one attempted way to overcome the lack of reliability of an observer (of the repetitive behaviours) was to introduce a second observer. This alone does not solve the problem, indeed it simply introduces another dimension for reliability problems. Candidates need to be aware that reliability issues must be addressed directly, for example through operational definitions or practice with recorded behaviours to develop a consistent recording ability within or between observers.

Another frequent 'solution' offered to overcome a lack of consistency between observers was to conduct a test of inter-observer reliability. This alone does nothing to overcome the problem, it merely demonstrates whether there is or is not a problem. This must be followed by a procedure to improve that consistency, such as elaborating on operational definitions of behavioural categories, practicing data collection using those definitions e.g. by working together with videoed behaviours and discussing and addressing differences.

Finally, a small number of responses described sampling issues, in contravention of the rubric, so could not earn marks.

© 2020

PSYCHOLOGY

Paper 9990/32

Paper 3 Specialist Options: Theory

Key messages

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

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

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

These questions could ask the candidate to describe a theory, self-report, or part of a study. These questions could also ask the candidate to describe a part of one of the named studies from the syllabus or a summary of the key features of the study. This question is worth four marks and the candidates should write a more extended answer. Some candidates did not provide enough detail of the theory, study or self-report.
A minority of candidates evaluated at the end of their description which was not creditable for the question set. Creating a summary of the theories, self-reports and studies from the syllabus could be useful for candidates to answer this type of question.

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

These questions could require the candidate to explain strengths or weaknesses of what they have described in the **part (b)** of the question. The question could also ask the candidates to make a comparison or to evaluate using a specific issue, such as practical applications. This question is worth six marks so the candidate should write a more extended answer for each issue raised. Some responses were very detailed for one issue but then only briefly discussed the second issue. In addition, some of the responses were general and not specific to the theory, self-report or study named in the question. To improve, responses should give specific examples to achieve the top band.

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

This question will always come from one of the bullet points in the syllabus. Candidates could describe the three or four studies, theories, self-reports, or techniques identified in the specification under the appropriate bullet point. For this exam, some of the answers did not give all of the studies/theories under the bullet point, used the incorrect bullet point or the description was brief. It is possible for the responses to achieve full marks by describing at least two of the studies, theories, self-reports or studies but this would need to be a very detailed description. Ideally the response would describe three of the bullet points in detail with excellent understanding and good use of terminology throughout. These types of responses often achieved the top band. It is also important that the descriptions are linked to the topic area named in the syllabus. It could be useful for candidates to do revision notes with the title of each bullet point as the header in their notes.

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

This question will always ask the candidate to evaluate the theories, studies and/or techniques described in **part (a)** of the question. The response must include at least two evaluation issues, including the named issue, in order to be considered to have presented a range of issues to achieve the top band. However, most responses that evaluated 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 at least three issues tended to

© 2020

achieve higher marks as these responses were able to demonstrate comprehensive understanding with good supporting examples from the theories, models, studies and self-reports described in the **part (a)** of the answer. The candidate must also provide some form of analysis. This could be done by discussing the strengths and weaknesses of the issue being considered, presenting a counter-argument to the issue under discussion or comparing the issue between two studies and/or theories. A conclusion at the end of each issue would be helpful in order to show excellent understanding of the issue under discussion. In order to achieve the requirements of the Level 3 and Level 4 band descriptors it would be preferable if the response was structured by issue rather than by study and/or theory. It would also be ideal for the response to start with the named issue to make sure the answer covers this requirement of the question.

Quite a few of the answers were structured by self-report/theory/study rather than by the issue which often led the response to be quite superficial and repetitive. A number of the responses did do analysis. Candidates should be aware this question is worth 10 marks and attempt to include an appropriate amount of information.

General comments

Candidates achieved marks across the full range of the mark band. Many of the candidates were very well prepared for the exam and showed good knowledge, understanding and evaluation throughout their responses. However, some candidates were not as well prepared and showed limited knowledge and understanding with brief and/or superficial responses. These candidates often had limited evaluation skills.

Time management for this paper was good for the vast majority of candidates and most attempted all questions that were required. Some candidates appeared to spend too long on the first option and left themselves less time to answer their second option. 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.

Comments on specific questions

Psychology and Abnormality

Question 1

(a) This question was answered well with the majority of responses focussing on biochemical causes and this was answered well on the whole. Some candidates confused two causes by, for example, linking positive reinforcement (behavioural) with feeling state (cognitive). Most responses were able to outline the cause of the impulse control disorder but did not then outline how this cause could lead to the behaviour becoming addictive or compulsive.

(b) Most responses achieved at some marks for this question by giving details of the Grant et al. study. Popular details included the sample, length of the study, details of the medication given to the participants and results of the study. However, many candidates were under the impression that Grant et al. gave the participants opiates rather than opiate antagonists.

(c) The vast majority of responses described at least one weakness of the Grant et al. study and many gave specific examples from the study to support their point. Common weaknesses mentioned included lack of generalisability to other impulse control disorders as the participants were only gamblers; the problems of using the Y-BOCS as a self-report as participants may not be truthful; the lack of follow-up so researchers do not know how long the effects lasted for; and the side-effects of the opiate antagonists. A common error was pointing out that participants were deceived if they were in the placebo group as this violated ethics. However, taking part in a double-blind drug trial the participants would be aware that they could either be given the opiate antagonist or the placebo and would have consented to this prior to the study.

© 2020

Question 2

(a) Many responses were detailed, accurate and coherent with a good use of psychological terminology. Most responses referred to the types of anxiety disorders and the measures including the BIPI and GAD-7. Some responses described case studies of anxiety disorders including Little Hans, little Albert and Kimya. A few described the Saavedra and Silverman (2002) study from the AS syllabus which was creditworthy. Characteristics of generalised anxiety were not always well described in the responses. A large proportion of responses simply listed names of phobias without distinguishing between agoraphobia and specific phobias. However, the BIPI and GAD-7 measures were generally detailed. Many responses did not describe the case studies in **part (a)** and this made it more challenging for these candidates to evaluate case studies in **part (b)**.

(b) The responses to this question covered the full range of the mark band. Stronger responses used the issues as a starting point and compared the types of anxiety disorders, measures and case studies that had been described in **part (a)**. Most addressed the named issue of case studies. Some did provide analysis of this issue and provided strength(s) and weakness(es) of case studies using the studies described in **part (a)** as examples. Weaker responses tended to describe case studies without providing any evaluation or evaluated the case studies very briefly at the end of the description. Some responses did provide strengths and weaknesses of case studies but did not give any examples and therefore the evaluation was limited. A range of other evaluation points were considered including self-reports, cultural bias, application to everyday life in terms of diagnosis and validity.

Psychology and Consumer Behaviour

Question 3

(a) Most responses were able to explain the 'disrupt-then-reframe' sales technique and used a clear example to illustrate their explanation. Most achieved full marks for this question. A few candidates used the words 'disrupt' and 'reframe' in their response without providing an explanation of what this meant which was not creditworthy.

(b) A full range of marks was given for responses to this question. Some candidates had a clear understanding of the theory of planned behaviour and often used a diagram for clarification. The vast majority applied this to purchase decisions. However, a significant minority of responses confused this theory with black box or (more rarely) consumer decision making model.

(c) For those candidates who gave a good response to **part (b)**, many were able to achieve level 2 and for their response to this question. Even where **part (b)** was answered well, very few were able to make suggestions that clearly linked to the theory of planned behaviour in order to achieve Level 3. Some good ideas included advertising where the popularity of a product was clearly shown (to alter subjective norms) and price cuts to increase perceived behavioural control.

Question 4

(a) Some responses gave a good outline of the studies in this bullet point including cognitive maps of retail locations (Mackay and Olshavsky, 1975), crowding in retail environments (Machleit et al., 2000) and shopper movement patterns (Gil et al., 2009). Many responses gave details of the sample, procedure and results of each study. The study by Mackay and Olshavsky appeared to be the least well understood of the three.

(b) Those responses that answered **part (a)** well tended to produce good answers to this question, with some good understanding of reliability and how it applies in the studies. However, few responses were able to produce any kind of analysis that gave the extent of an issue or debate or examined the strengths and/or weaknesses of an issue they were using. Weaker answers tended to list five or six issues and/or debates and provided a very brief application. A significant number of responses evaluated the lack of ethics in the study by Gil et al. on shopper movement patterns although participants were aware of the observation of their movement and why they had a coloured tag and thus had given consent. Popular evaluation issues included validity, reliability, generalisability, practical applications and ethics.

© 2020

Psychology and Health

Question 5

(a) There were many good responses to this question explaining that the tokens/stamps were given for avoiding lost-time injury and these tokens/stamps could be exchanged for items in a store. Some of the responses were overly long for a two-mark question. In order to give appropriate time to responses, candidates should be advised of the importance of being concise.

(b) There were some strong answers to this question, demonstrating sound knowledge of the Tapper et al. 'whole-school' food dude program, summarising the materials used, use of stickers to reward trying a new food, how rewards became more intermittent and the gathering of data from home. Weaker responses included fewer details of the program and sometimes included results which were not creditworthy.

(c) Responses to this question were usually able to provide a strength of the study by Tapper et al. including large sample size, practical applications of increasing fruit and vegetable consumption in children and high ecological validity. However, a large proportion mistakenly described the program as unethical when consent had been obtained from parents and teachers and ultimately the program should improve children's health. Another common weakness given was that the study lacked generalisability because it was only carried out in one school, whereas this was not the case as a number of schools from a variety of areas of the UK were involved. However, other candidates correctly pointed to the study being ethnocentric as it only took place in UK and may not apply in other cultures, somewhat expensive to use (potentially also leading to a cultural bias) and the self-report portion being open to social desirability bias.

Question 6

(a) This was generally a well answered question where candidates demonstrated sound knowledge of hypochondriasis, Munchausen syndrome and (to a lesser extent) the study by Safer et al. of delays in seeking treatment. Some responses were very detailed and could achieve in the higher mark bands. Weaker responses gave superficial descriptions of Munchausen syndrome, hypochondriasis and some detail of the types of delays in seeking treatment investigated by Safer et al. Some responses were from the incorrect bullet point and gave descriptions of adherence (pill counting) or health belief model which was not creditworthy.

(b) A number of responses structured their answer by addressing each issue in turn. Most responses considered the named issue of practical applications and discussed how this could provide a better diagnosis to target sufferers of Munchausen syndrome and hypochondriasis but often did not explain how this could be done or how it might help these patients. Other popular issues included generalisability, validity and self-reports. Some responses achieved in the lower levels of the mark band due to giving very brief responses or structuring their response by study/type of disorder which meant these types of answers were often repetitive and superficial.

Psychology and Organisations

Question 7

(a) There were some very good responses to this question and most wrote an appropriate amount for a two-mark question. Most responses could identify what is meant by job satisfaction. There were some good examples of where satisfaction can be found in motivation, pay or conditions. However, some responses used the words in the question in their answer, saying, for example, that job satisfaction is when someone is satisfied with their job, which does not demonstrate understanding. Candidates should be encouraged to use their own words to avoid this.

(b) Most responses were able to give a reasonable description of the quality of working life (QWL) questionnaire by Walton. Many responses stated that it assessed eight key components and were able to list the components measured. A few responses also described the Likert scale that was used for the responses to this questionnaire.

© 2020

(c) Examples of points correctly made in this question were the problem of social desirability bias leading to employees saying they were satisfied with their job when they were not and the use of quantitative data leading to easy comparisons. However, few responses were able to contextualise their answers and tended to provide rather generic responses to self-reports rather than job satisfaction measures. Many responses lacked structure.

Question 8

(a) There were some good, well developed responses to this question. Many responses described universalist and behavioural theories, adaptive leadership (Heifetz, 1997) and the three levels of leadership (Scouller, 2011). Weaker credit-worthy responses tended to be brief or a superficial description of the relevant theories. Some responses were anecdotal with the response describing what would make a good leader.

(b) Many responses were structured by evaluation issue with many of them beginning with the named issue of nature versus nurture. Some responses did do some analysis of their evaluation points by making a comparison between the theories that had been described in **part (a)**. Popular evaluation issues included applications to everyday life and generalisability/cultural bias. A significant number of weaker responses evaluated the theories from **part (a)** in turn and gave more superficial and repetitive responses. Little analysis was given as candidates tended to identify if the theory supported either the nature or nurture side of the debate with little explanation given for this or why it is helpful to evaluate this issue. In addition, some theories may lie in between these two extremes.

© 2020

PSYCHOLOGY

Paper 9990/42
Paper 4 Specialist Options: Application

Key messages

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

General comments

Section A

- Candidates did not always address the 'stem' of the question, the introduction or the opening words in Section A when this is crucial to answering each question part that follows.
- Answers must refer to the study the question is about. Many answers made general comments which did not demonstrate their understanding about the study itself (see specific questions below for examples).
- Many answers correctly included advantages and disadvantages but many needed to relate these to the question to access full marks.
- In part (d) questions, many conclusions repeated what had already been written, and such summaries could not be credited. A conclusion is a 'decision reached by reasoning' and so as the reasoning has been done through the advantages and disadvantages, a final decision/conclusion needs to be drawn.
- Candidates are required to consider what the question requires rather than writing pre-prepared answers. Many questions will test the ability to apply knowledge from one thing to another, particularly methodological knowledge.
- Candidates should always provide sufficient detail to score all the available marks. A single sentence is more likely to score 1 mark rather than 2 marks, so a little elaboration, explanation or example that goes beyond the basic sentence is always recommended. Candidates should always try to impress the Examiner with their psychological knowledge.

Section B

Answers to **part (a)** questions in this section should include an appropriate design, have applied a range (four or five) relevant methodological design features, each of which should be explained fully, showing good understanding. Many answers listed features such as 'I would have a random sample' and 'It would be an independent measures design' without explanation of why it would be a random sample, or how this would be obtained.

© 2020

In **part (b)**, answers should explain the methodological decisions on which their **part (a)** design is based and also explain the psychological evidence on which their design is based. Merely describing a relevant piece of research from the topic area is insufficient and cannot be credited. The links between the research and how it informed the design must be shown. Some responses included statements such as 'I chose a self-selecting sample because Milgram (1963) did', which identified a study using that technique but did not explain the choice of sampling technique.

Section C

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

Comments on specific questions

Section A

Question 1

(a) Many candidates scored full marks by identifying 'sad', 'crying' and 'loss of weight'. Some candidates scored 1 mark for one correct feature and some candidates scored no marks by appearing to guess and stating items that were not on the BDI.

(b) Many candidates answered the question in detail and scored full marks, with answers explaining how depression resulted from the cognitive triad and negative automatic thoughts. However, many candidates did not answer the question, instead focusing on the BDI itself. Such answers described the rating scale, scoring and other features. The question focused on the theory on which the BDI is based, and the theory is the former type of answer rather than the latter.

(c) (i) The reliability of a questionnaire, such as the BDI, can be assessed using test-retest (by administering the *same test* to the *same person* on two different occasions and comparing the result) and candidates writing about test-retest scored full marks. Split half reliability was also creditworthy. Many candidates were not able to demonstrate their understanding of reliability, or they confused it with validity. Some candidates were able to demonstrate their understanding of reliability, but not how reliability can be tested. Other candidates incorrectly suggested that inter-rater reliability can be used, which was not applicable here.

(ii) The way to test the BDI is to use concurrent validity (where the BDI is compared to an existing measure of depression). Candidates writing about concurrent validity scored full marks. Some types of validity do not apply here, such as population validity and ecological validity. Many candidates were not able to demonstrate their understanding of validity.

(d) Many answers included two advantages and two disadvantages and a conclusion, but only scored partial marks because often measuring depression as the question required was not addressed. Candidates appeared to have pre-prepared answers for a question on self-report questionnaires. This strategy is not recommended because generic answers cannot achieve full credit, for which the question must be addressed in full. Most candidates gave the advantage that questionnaires produce objective data, and gave the contradicting disadvantage that questionnaires are flawed because of social desirability, response bias, etc. Questionnaires are always subjective and can never be objective because they are self-reports.

Question 2

(a) Many answers scored full marks because they explained what the term multiple unit pricing meant and they gave an example, often using the example provided in the stem ($.5 per can or 4 for $2). Some candidates misunderstood the term thinking that there was a discount involved. There is not. It is a 'psychological trick' to make the consumer think they are getting a bargain when they are not.

© 2020

(b) Many candidates wrote about sales techniques that were not 'point of purchase' techniques, as the question stated, which could not be credited. The two main point of purchase techniques are purchase quantity limits and suggestive selling. Marks were awarded for identifying the basics of these terms, but full marks were only awarded if examples were provided as required by the question. Some candidates answered the question fully and received full marks. An example of purchase quantity limits is to advertise 'limit of six cans per customer' and an example of suggestive selling is 'buy ten for your freezer'.

(c) Most candidates scored 2 marks for providing one strength and one weakness. However, answers were usually general and could apply to any question. Some candidates did not answer the question fully. This question stated 'students as participants … in research on purchase quantity decisions' meaning that the strength and weakness had to be related to this specific topic area and providing an example would have been sufficient to earn marks.

(d) Many answers included two advantages and two disadvantages but often focused exclusively on students without addressing purchase quantity decisions. To score full marks, advantages and disadvantages must be related to the study or topic in question. Sometimes relevant conclusions were provided, but often a summary was provided instead which could not be credited.

Question 3

(a) A psychometric test is a 'measure of the mind' or in this instance a 'measure of pain' and relevant examples would include the McGill Pain Questionnaire (MPQ) or any similar test such as the PPQ. Most answers scored full marks by including these two components. A clinical interview is not psychometric, but such answers did score a partial mark if the term 'psychometric' was explained.

(b) Nearly all candidates scored full marks by explaining the difference between the terms (acute is short-term, chronic is long-term) and by providing examples of each (acute such as finger trapped in door, chronic such as arthritis). A small number of candidates confused the two terms.

(c) Very few answers scored full marks, because candidates did not always answer the question set. The question stated 'use an example in your answer'. Most candidates scored limited marks for giving one advantage and one disadvantage. Some candidates gave several of each, but only one of each could be credited. A few candidates did not appear to understand the term 'clinical interview'.

(d) This question required advantages and disadvantages of using psychometric tests to measure pain. Many good answers provided both advantages and disadvantages, but only a few candidates went on to support each advantage/disadvantage with an example. Many candidates summarised their answer at the end, instead of providing a conclusion, which could not be credited.

Question 4

(a) Many candidates did not score marks because they did not demonstrate their understanding of the term, e.g. 'an open plan office is an office that has an open plan'. Others described the image provided in the stem of the question, but this was insufficient for full credit. A psychologically informed explanation was required. An open plan office places worker desks in large, open spaces, to create the feeling of space and minimises the use of small, enclosed rooms such as private offices.

(b) Many candidates appeared not to know any job characteristics, such as those by Hackman and Oldham, instead focusing on what might be different about working in an open plan office, which could not be credited. Strong responses covered about autonomy, task identity, supervisor and co-worker feedback or friendship opportunities. Occasionally skill variety, task significance and task feedback were mentioned, which also received credit.

(c) Two dependent variables were required. Many candidates wrote about office design affecting productivity, despite this not being a variable measured by Oldham and Brass. Relevant answers included: work satisfaction: the degree to which an employee is satisfied and happy with the job; interpersonal satisfaction which is satisfaction with co-workers and supervisors; and internal work motivation which is how a worker feels about the job. Whilst many candidates scored full marks, others scored no marks.

© 2020

(d) The question required a discussion of interviews in relation to open plan offices. There were some excellent answers which followed a formula (which could apply to any question **part (d)** in this **Section A**): advantage plus example; advantage plus example. Disadvantage plus example; disadvantage plus example. Conclusion (not a summary).

Section B

Question 5

(a) Candidates could score more marks if they thought through their design before starting their response. Many candidates opted to conduct a laboratory experiment when ECT should only be conducted in a hospital, and by a medically qualified practitioner. There was the incorrect assumption that a longitudinal study is automatically a case study of one person. It is not; longitudinal studies can be conducted on large samples. Many candidates suggested designs that did not test for side effects as the question required. Some candidates proposed giving a BDI before and after treatment to see if the patient had improved, despite the BDI not measuring side effects. A final problem was that candidates would administer ECT over time, rather than record the side effects that might develop over time.

(b) Many candidates for their psychological evidence wrote about ECT in general, mentioning how it was discovered and how it is said to work, for example. What was frequently absent was a focus on side effects, such as memory loss, as required by the question. The psychological evidence included in this question part should focus specifically on explaining the basis of the **part (a)** design meaning that if a study has not informed the design in **part (a)** then it should not be included in **part (b)**. In relation to methodological evidence, candidates could have explained how they would measure any side effects over time, logically by using a questionnaire repeated periodically over time. This aspect would focus specifically on explaining a design decision directly relevant to the question asked.

Question 6

(a) There were many good answers, with candidates inventing competent designs. The common features of IV (usually high- versus low self-monitors), DV, controls and experimental design were often included although they were not always explained in very much detail. Answers at the top end of the mark range included general factors such as sampling technique and the type of data gathered. Rather than writing about multiple factors in one sentence, a list-like response, more marks can be earned by reducing the range and adding more depth to explanations.

(b) Psychological knowledge in the stronger answers showed a good understanding of the study by Snyder and DeBono because answers in **part (a)** were often informed by this study. Supporting evidence was less strong on brand recognition, although some candidates did base their designs on the work of Fischer. In relation to methodological decisions, answers were generally coherent because candidates knew the features of an experiment. However, answers were often superficial. If an independent measures design is used, then the reasons for its choice over a repeated measures design should be given, rather than just state ‘I used an independent design’.

Question 7

(a) This question required candidates to use a questionnaire. Many of the basics of questionnaire design were included, but often these were either not thought through or were not related to rational non-adherence. For example, some questionnaires began with ‘what is your name?’ and are you ‘male/female?’ and then later in the answer it was claimed that the study was ethical despite breaking the confidentiality guideline. Questions should have been based on the work of Bulpitt and rational adherence, but often questions were general, such as ‘why do you not adhere’. Answers must be focused on the topic area in question and answers should include all the relevant methodological features.

(b) In relation to methodology, candidates often stated that they gathered quantitative data. However, there was rarely a mention of why this was an advantage or there was no discussion about why qualitative data was not gathered. Also unclear in many designs was where and how the questionnaire was conducted. Was it a postal questionnaire, or presented face-to-face, for example. Using a postal questionnaire would have brought in other studies where this was done, such as the study by Riekart and Droter, whose response rate was low and so face-to-face would

© 2020

be a better choice. Discussion of a few points explaining the design in detail gains marks. Psychological knowledge should have focused on rational non-adherence. In many instances it did, but not in every answer.

Question 8

(a) Candidates had a free choice of method here, and the strongest answers chose an experiment, so metropolitan and continental rotations could be compared. Some candidates gathered data using a questionnaire, but often the design of the questionnaire was lacking. The questions asked should have been based on knowledge of the two rotations, rather than just 'which shift do you like best'. This is what links this **part (a)** answer with **part (b)**. Some candidates provided what they claimed to be a hypothesis, but often this was just a statement about what they expected to happen. A hypothesis, following on from knowledge acquired at AS, should be appropriately tailed with the IV and DV clearly included.

(b) The psychological evidence included in this question part should focus specifically on explaining the basis of the **part (a)** design. For this question both the metropolitan and continental rotation systems were often described, but often answers did not explain how they related to **part (a)** of this question. As written in **part (a)** if data is gathered by questionnaire, for example, then knowledge about both rotations should inform the question that are asked.

Section C

Question 9

Many candidates failed to score high marks because they did not answer the question set. The focus of the question was on cognitive explanations of schizophrenia and the nature versus nurture debate. There were three common types of answer. First, those which described the cognitive explanation, that by Frith, and did not refer to nature or nurture. Next, there were answers that described the cognitive explanation and then evaluated it with points about reductionism and similar, briefly mentioning nature or nurture. Finally, there were those answers who began by stating the nature versus nurture debate and then presented evidence to support one side or the other, with the cognitive explanation featuring as evidence. This latter type of answer scored the most marks of the three types.

Question 10

There were some good answers in response to this question showing that candidates knew about both generalisations and the study by Hall et al. There were many 'India-centred' answers, suggesting that people in India know about tea (used in the design of the Hall et al. study) whereas in other countries they know less, and hence generalisations are problematic. This was a good point and received credit. This point seemed to stimulate a wide range of other relevant points and some very good answers were written.

Question 11

Like other answers for different options in this Section many candidates described what they knew, rather than organising and using information to address the question. For example, there were some descriptive essays on adherence and this type of answer scored no marks because it does not answer the question set. This was exacerbated because the question focused on practitioner style and adherence, not just adherence.

Question 12

Three common types of answer were evident. Firstly, many candidates described Maslow's hierarchy of needs, often in detail. However, this is not a 'describe' question. The question asks for a discussion of whether the hierarchy is culturally biased or not. Secondly, candidates described and then evaluated Maslow's theory, with cultural bias often not mentioned. Thirdly, there were answers which answered the question, presenting arguments both for and against cultural bias and using examples from Maslow's theory in support. This latter type of answer scored much higher marks; the former two types scoring no more than bottom band marks.

© 2020