Pearson
Edexcel

Examiners' Report
June 2022

GCE Psychology 9PS0 03

ResultsPlus

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications come from Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk.

Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

ResultsPlus

Giving you insight to inform next steps

ResultsPlus is Pearson's free online service giving instant and detailed analysis of your students' exam results.

- See students' scores for every exam question.
- Understand how your students' performance compares with class and national averages.
- Identify potential topics, skills and types of question where students may need to develop their learning further.

For more information on ResultsPlus, or to log in, visit www.edexcel.com/resultsplus. Your exams officer will be able to set up your ResultsPlus account in minutes via Edexcel Online.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk.

June 2022

Publications Code 9PS0_03_2206_ER

Introduction

The summer 2022 examination was the first full exam series since 2019 and was taken by candidates affected by the pandemic. As such, advance information was provided to help candidates to focus their revision time and offered them details of the focus of the content in the 9PS0/03 examination. The performance of the candidates is summarised below, with advice on how to improve in future series.

As with previous sittings of this paper, the best responses in section A gave fully contextualised responses related to the novel scenario given in the question. Some candidates continue to produce generic responses, and some also found difficulty with some aspects of research methods, including cross-sectional designs, interpreting the histogram, and standard deviation. As such, there are some areas that centres can focus on to help support candidates for future series. It may be that some candidates focused so much of their time on section B and C, perhaps due to the advance information provided in which section A could not be included, that they placed less emphasis on this section than usual. It was also noticeable that the skills required to answer some of the questions in this section were lacking within candidate responses.

Regarding section B, performance was mixed. Emphasis was clearly placed on Q4 by candidates, perhaps due to the advance information provided. As such, performance in Q3 was varied with some very good responses, but also some very weak responses or non-attempts. It was clear that candidates were more prepared for Q4 as there was a higher standard in general than previous series.

Section C also produced mixed performance from candidates. Similar to Q4, candidates were clearly more prepared for Q6, perhaps due to the advance information provided. As such, this did produce a generally higher performance than previous series. Question 5 had a lower performance in general than previous series, despite the question using the same style of assessment as previous series and advance information being provided. This may have been due to candidates focusing so much of their time on Q4 and Q6, which had noticeably more content in general than previous series.

The remainder of this Examiner Report will focus on each individual question and specific examples of candidate responses which can be used to help prepare students for future 9PS0/03 examinations.

Question 1 (a)

Question 1a required candidates to explain a strength and weakness of using a cross-sectional design in the study. The best responses identified a strength and weakness of using a cross-sectional design in context and then fully justified their points. Weaker responses identified a strength and weakness but did not fully justify their points, or gave generic content or inaccurate ideas. Candidates generally found this question difficult, with only the minority achieving the higher marks in the range, so this should be a focus for centres supporting candidates for future series.

(a) Explain **one** strength and **one** weakness of using a cross-sectional design for the junior to senior Pinocchio study.

(4)

Strength

cross-sectional design can be used to investigate a large group of participants in one moment to make comparisons. For example, in the Junior to Senior Pinocchio study researchers are able to investigate people of different age groups regarding their lying ability/ willingness in one moment, making it more time efficient and quick.

Weakness

A weakness of cross-sectional design is that you cannot investigate in detail a particular group. For example, unlike, a longitudinal study researchers are unable to investigate further into why a certain group lie more or less as the cross sectional design only study a range of different participants.

ResultsPlus
Examiner Comments

This response was awarded 2 marks.

Two marks for the strength – one for identification of the strength and one for justification.

No marks for the weakness – this is not rewardable.

Question 1 (b)

Question 1b required candidates to explain two conclusions using the data from the table. The best responses identified two relevant conclusions and then justified each conclusion using evidence from the table. Weaker responses gave the conclusions only, or interpreted the data inappropriately so gave inaccurate conclusions or just recycled the data from the table with no conclusions presented. Performance on this question was mixed, with candidates most commonly achieving 4,2, or no marks, with the most common being four marks.

(b) Explain **two** conclusions you can make using the data in **Table 2** regarding the lying ability of the participants in the age categories.

(4)

1 Participants in young adulthood (18-29) ~~have the~~ best are the most able liars. This is because the average response time of 3.25 is much lower than any other age category and ~~error~~ error rate of 8.55 is also the lowest suggesting that they don't find it difficult to lie.

2 Participants in early childhood (age 6-8 years) are the least able liars. This is because the average response time (3.83) and error rate (20.20) ~~was~~ is much greater than all other age groups suggesting that younger children find lying more difficult.

ResultsPlus
Examiner Comments

This response was awarded 4 marks.

First conclusion – one mark for identification and one for justification through analysis/interpretation.

Second conclusion – one mark for identification and one for justification through analysis/interpretation.

Question 1 (c)

Question 1c required candidates to explain a weakness of the study in terms of validity. The best responses identified a weakness of the study in terms of validity and then justified the weakness given. Weaker responses identified a weakness only, gave a weakness in terms of something other than validity or gave a generic response. The most common responses focused on either the subjective nature of the research or how realistic the task was considered. Candidate performance was split fairly evenly across the mark range, so performance was varied.

(c) Explain **one** weakness of the junior to senior Pinocchio study in terms of validity.

(2)

This study has low ecological validity. This is because the lying task the participants performed - pressing a button to answer a question depending on its colour - does not reflect how or why people lie in real life, usually in conversations and because of embarassment for example. So, the results are not true to lying how lying works in real life (ecologically invalid)

ResultsPlus
Examiner Comments

This response was awarded 2 marks.

One mark for identification of the weakness, one for justification.

Question 1 (d)

Question 1d required candidates to explain a strength of the study in terms of reliability. The best responses identified a strength of the study in terms of reliability and then justified the strength given. Weaker responses identified a strength only, gave a strength in terms of something other than reliability or gave a generic response. The most common response was to focus on the standardised questions given by the researchers. Similar to Q1c, candidate performance was split fairly evenly across the mark range, so performance was varied.

(d) Explain **one** strength of the junior to senior Pinocchio study in terms of reliability.

(2)

The study has high ~~internal~~ reliability, as a standardised procedure was used. This is because all participants responded to the same 15 questions, and colours were used in similar ways to indicate a yes/no response and induce lying. So, as the conditions remained the same except for age, the lying was measured in a consistent fashion, giving the study reliability.

ResultsPlus
Examiner Comments

This response was awarded 2 marks.

One mark for identification of the strength, one for justification.

Question 1 (e)

Question 1e required candidates to state an appropriate conclusion with reference to the type of skew shown in the figure. The best responses clearly identified the skew and gave an appropriate conclusion. Weaker responses did not refer to the type of skew as specified in the question, or gave an inappropriate conclusion, with some also misinterpreting what the histogram was showing. Candidates generally found it difficult to identify the correct skew and give a conclusion, so this should be a focus for centres to help support candidates for future series.

(e) State **one** conclusion that could be made from **Figure 1** with reference to the type of skew shown.

(1)

The g histogram presents itself with a positive skew, which suggests that a higher percentage of participants barely lied within the last 24 hours (0-1)

ResultsPlus
Examiner Comments

This response was awarded 1 mark.

One mark for an appropriate conclusion with reference to the skew.

Question 1 (f)

Question 1f required candidates to calculate the mean score for the data to one decimal place. The vast majority of candidates were able to give the correct mean score to one decimal place.

(f) Calculate the overall mean lying frequency for the participants using the data in **Table 3**. You must give your answer to **one** decimal place.

(1)

SPACE FOR CALCULATIONS

$\frac{1.75 + 2.5 + 3 + 2.5 + 2 + 1.75 + 1.5}{7} = 2.1$

Mean 2.1

ResultsPlus
Examiner Comments

This response was awarded 1 mark.

One mark for the mean to one decimal place.

Question 1 (g)

Question 1g required candidates to explain a weakness with using the volunteer sampling technique for the study. The best responses identified a weakness of using a volunteer sampling technique in context and then justified the weakness given. Weaker responses identified the weakness only, gave a generic response, or gave inaccurate content. There were more generic responses for this than other questions, so this should be a focus still for centres supporting candidates for future series. Overall, performance was varied with those who gained marks, similarly split across one and two marks.

(g) The researchers used a volunteer sampling technique to gather the participants for the junior to senior Pinocchio study.

Explain **one** weakness with using a volunteer sampling technique for the junior to senior Pinocchio study.

(2)

People that volunteer to take part in a study to test lying ability may have a better lying ability than the average person, causing them to take part. This means that results regarding lying ability may not actually be representative of the wider population's lying abilities

(Total for Question 1 = 16 marks)

ResultsPlus
Examiner Comments

This response was awarded 2 marks.

One mark for identification of the weakness, one for justification.

Question 2 (a)

Question 2a required candidates to calculate the standard deviation of the data to two decimal places. The best responses showed all their working and gave the correct response to two decimal places. The majority of candidates found this difficult though and whilst a lot attempted the calculation, they did not achieve any marks. Standard deviation should remain a focus for centres supporting candidates for future series.

(a) Calculate the standard deviation for the vocabulary score using the data in **Table 4**. Show your working and give your answer to **two** decimal places.

(4)

SPACE FOR CALCULATIONS

$\sqrt{\left(\frac{\Sigma(x-\bar{x})^2}{n-1}\right)}$

$n = 9$

$n-1 = 8$

$\bar{x} = \frac{900}{9} = 100$

x	$x-\bar{x}$	$(x-\bar{x})^2$
100	0	0
87	-13	169
105	5	25
92	-8	64
98	-2	4
107	7	49
101	1	1
96	-4	16
114	14	196

$\Sigma(x-\bar{x})^2 = 524$

$\sqrt{\frac{524}{8}} = \sqrt{\frac{131}{2}}$

$= 8.09$ (to 2dp)

Standard deviation 8.09

ResultsPlus
Examiner Comments

This response was awarded 4 marks.

Four marks for the standard deviation to two decimal places.

Question 2 (b)

Question 2b required candidates to explain a reason for using the standard deviation rather than the range. The best responses identified a suitable reason for using the standard deviation rather than the range and then justified the reason given. Candidates generally found this difficult with a lot stating that the standard deviation is not affected by outliers / extreme scores, which is inaccurate so achieved no marks. As such, supporting candidates to understand why the standard deviation may be more useful than the range should be focus for centres for future series.

(b) Explain **one** reason for using the standard deviation rather than the range as a measure of dispersion.

(2)

Standard deviation is less likely to be effected by extreme values but range is more likely to be effected by extreme values.

ResultsPlus
Examiner Comments

This response was awarded 1 mark.

One for identification of a suitable reason only.

Question 2 (c)

Question 2c required candidates to explain an improvement for the study. The best responses identified an appropriate improvement in context and then justified the improvement given. Weaker responses identified an improvement only, gave an inappropriate suggestion or gave a generic response. Performance was mixed, with a spread of marks being achieved by candidates, but generally they found it difficult to give a fully justified improvement in context.

(c) Explain **one** improvement that could be made to the 'textisms' and literacy study.

(2)

One improvement could be to use children from different countries, this could be done by asking for children volunteers from the USA, Germany, France, to and would increase the representativeness of the sample of children and make it generalisable to all children age 10 to 12's level of vocabulary and + extisms score.

(Total for Question 2 = 8 marks)

ResultsPlus
Examiner Comments

This response was awarded 2 marks.

One for identification of an appropriate improvement, and one for justification.

Question 3 (a)

Question 3a required candidates to compare the observed/calculated value with a relevant critical value and then justify what this means for the study. The best responses gave a focused response which included both elements, whereas weaker responses tended to focus on one or the other or focus on the data from the table instead. Generally, candidates found it difficult to provide both a comparison of the critical and calculated values and then interpret this in terms of the study. Some candidates misunderstood the study and so their interpretation was not creditworthy.

(a) The researchers wanted to see if there was a significant difference in the perceived effectiveness of crying. They conducted a Mann-Whitney U test and found an observed / calculated value of 110 for a 5% level of significance with a two-tailed test.
calc & c.v
C.V = 112
Explain what this shows in terms of the perceived effectiveness of crying in the reconciliation after romantic conflict study.

(2)

The calculated value of 110 is lower than the critical value of 112, this shows a significant difference in the perceived effectiveness of crying ~~for~~ by males and femaus in the reconciliation after romantic conflict study.

ResultsPlus
Examiner Comments

This response was awarded 2 marks.

One mark for comparing the observed/calculated value with a relevant critical value, and one for justification of what this means for the study.

Question 3 (b)

Question 3b required candidates to explain how far social learning theory could account for the findings of the study, using research evidence. The best responses applied social learning theory to the findings of the study and then justified their ideas using research evidence. Weaker responses tended to give a lot of information about social learning theory and then not apply it to the study appropriately, but instead say that the participants reproduced the behaviour themselves. Candidates found it difficult to reach the higher marks in the range, with the majority focusing on application only and perhaps considering a single piece of research evidence, usually Bandura et al.'s (1961) study. The minority were able to apply various parts of the theory to the study findings and then provide various research to support or oppose their ideas as appropriate, but those reaching the top of the mark range were very infrequent.

(b) Using research evidence, explain how far social learning theory could account for the findings of the reconciliation after romantic conflict study.

(6)

Social learning learning could explain ~~how~~ why 'communicate' has the highest effectiverate per both male (5.44) and female (6.13). The couple would pay attention to ~~their~~ partners on TV trying to communicate ~~with them~~ to solve an argument between the TV couple. They would retain how ~~their partner~~ communicated, and then reproduce the behaviour by communicating in the same way as on TV back to their partner. The would be motivated to continue communicated as their partner would like the TV couple did. have a positive response. This is supported by Bandura's bobo doll experiment where he found that children imitate role models aggression. In the film-mediated variation the filmed group displayed as mean aggressive acts compared to the control grup which had 19 mean aggressive act.

ResultsPlus
Examiner Comments

This response was awarded 2 marks.

One mark for application of social learning theory to the findings of the study, and one mark for judgement/justification of research evidence in relation to the study.

Question 4

Question 4 was an extended open response question with the 'Evaluate' taxonomy which targets both AO1 and AO3 content. AO1 was looking for knowledge and understanding of the studies or ethical issues and AO3 was for analysis, interpretation, and evaluation of both studies in terms of how ethical they could be considered and the implications of this, leading to judgements/conclusions.

Assessment of this question was through a levels-based mark scheme where a 'best-fit' approach was used; deciding which level most closely describes the quality of the answer. Each AO was judged separately and where the components met the requirement for the level fully (and perhaps has elements of the level above), then marks were awarded at the top of the level. Where the components met the level but only just, they were awarded marks at the bottom end of the level. When a response was imbalanced (i.e. one AO was stronger than the other) a compromise was found. Consideration was also given regarding this question requiring greater AO3 content than AO1 (6/16 to AO1, 10/16 to AO3).

Performance was noticeably better on this question than in previous series, perhaps due to the advance information provided and the nature of the debate being assessed in relation to the studies. Candidates typically gave a lot of information compared to previous series and focused on both studies and considered the ethical issues for each. The best responses gave accurate knowledge and understanding of the two classic studies and then placed greater emphasis on the AO3 content in terms of considering how ethical they were and the pros and cons of this. Weaker responses gave vague, brief information regarding the studies and tended to give a lot of inaccurate statements regarding the ethics of the studies, sometimes interweaved with snippets of accurate information. Those achieving the highest marks imbalanced their response, with more AO3 than AO1 and depth in terms of their arguments concerning ethics for the two classic studies.

① PFH ✗ ④ AFH ✗
② IC ✗ ⑤ RTW ✗
③ RTW ✓ ⑥ Confidentiality ✓

4 Evaluate Watson and Rayner (1920) and Sherif et al. (1954/1961) in terms of ethical issues. **(16)**

Sherif et al. wished to study the effect of competition on levels of hostility between two groups at a summer camp in Robbers cave national park. An ethical issue with this is that the levels of conflict between the two groups of boys became dangerous, with fights breaking out, name calling and burning eachothers flags. Researchers also gave out knives as prizes for competitions, instigated violence themselves by raiding a groups cabin and were reluctant to intervene in the hostility as they wanted to study it. This to put participants in great harm and could have caused severe distress, thus effecting the credibility of the study.

Participants consisted of 20 Protestant, middle-class 11 year old white boys where parents gave permission for them to attend a Summer camp. An issue with this is that researchers did not fully inform the parents of the study and asked them to stay away, they also did not disclose any details of it to the participants. This is deception and breaches ethical guidelines as researches did not gain informed consent. However, it may be argued that the benefits outweigh the ethical costs as, by

being able to study naturally occuring prejudiced behaviour, the researchers can make valid conclusions that can be applied to the real world to reduce prejudice, such as jigsaw classrooms.

Sherif et al's study saw that two boys from the 'Eagles' group withdrew due to homesickness. This means that the study adhered to BPS Ethical and Conduct guidelines by allowing ~~partic~~ participants the right to withdraw from the study at any time for any reason ~~reason~~ making it more ethical and thus legitimate. However, by adhering to this guideline, researchers eneded the 300 hour matching process of the boys' sporting ability, behaviours and IQ as the groups were unequal and thus unfairly matched in competitions. The 2 Eagles leaving could have therefore skewed results so perhaps ethical guidelines can be justifiably breached.

Watson and Rayner's study aimed to investigate if a 9 month old boy could be classically conditioned to fear rats. This directly goes against protection from psychological or physical harm as they ~~attempt~~ struck a steel bar beyond the boy's Little Albert's head to elicit a fear response. This evidently caused

distress and they successfully conditioned him to fear white rats, which may have interfered ~~the~~ with his quality of life.

This ~~fear~~ breach of ethical guidelines may have been remedied by researchers via deconditioning if Little Albert's mother had not withdrawn him from the study before researchers were able to do so. However, this meant meant that they did adhere to the guideline of right to withdraw ~~which but~~ but this may have been ultimately detrimental to both Little Albert and the study. Researchers were not able to see if his fear could be deconditioned and he may have had to live the rest of his life with a phobia that was intentionally given to him. This negatively affects the reputation of psychology. But forcing participants to remain in the study takes away free will and may have had even worse consequence.

Finally, researchers adhered to the guideline of ~~and~~ confidentiality, where participants are kept anonymous to protect their real identity. This did this by using ~~a~~ the pseudonym 'Little Albert' as well as finding who his mother was. This was a positive thing as they show respect for the family and their identity, this is especially important as the

study so highly controversial and the family may have received hate for participating as well as prematurely withdrawing. Thus, ~~at~~ confidentiality is a good ~~thing~~ thing. However, by keeping Little Alberts identity a secret, researchers were not able to follow up long term to see if the conditioning remained or if Little Albert may have been abnormal in some way. This reduces test-retest reliability as well as validity of the study so ethical discrepancies as well as adherences interfere.

In conclusion, whilst both Sherif and Watson and Rayner caused harm to their participants, ~~it~~ they both largely adhered to guidelines and it can be argued any breaches ~~were~~ of them were justified as the findings of both studies benefited society, such as reducing prejudice with the Jigsaw classroom as well as treating phobias with classical conditioning with systematic desensitisation.

ResultsPlus
Examiner Comments

This response was awarded Level 3 – 11 marks.

The AO1 was judged as level 3 – Demonstrates accurate knowledge and understanding.

The AO3 was judged as level 3 – Arguments developed using mostly coherent chains of reasoning leading to a conclusion being presented. Candidate demonstrates a grasp of competing arguments but evaluation is imbalanced.

Question 5

Question 5 was an extended open response question with the 'Evaluate' taxonomy with a scenario which targets AO1, AO2 and AO3 content. AO1 was looking for knowledge and understanding of biological psychology, AO2 was application to the scenario given in the question about the ability to drive, and AO3 was analysis, interpretation, and evaluation of the ideas presented from biological psychology or how other alternative ideas can account for human behaviour and the implications of this, leading to judgements/conclusions.

Assessment of this question was through a levels-based mark scheme where a 'best-fit' approach was used; deciding which level most closely describes the quality of the answer. Each AO was judged separately and where the components met the requirement for the level fully (and perhaps has elements of the level above), then marks were awarded at the top of the level. Where the components met the level but only just, they were awarded marks at the bottom end of the level. When a response was imbalanced (i.e. one AO was stronger than the other) a compromise was found. Consideration was also given regarding this question requiring equal amounts of AO1, AO2, AO3 (4/12 to AO1, 4/12 to AO2, 4/12 to AO3).

General performance on this question was lower than in previous series as there were more blank responses. This was perhaps due to candidates using more of their time on Q4 and Q6, which had noticeably more content than previous series in general. The best responses gave a balanced response with consideration of biological theory, application to the scenario, and then analysis, interpretation and evaluation leading to judgements/conclusions. The most common ideas that candidates focused on included hormones, genes, brain structure and function, with some also giving consideration to Freud's ideas. Weaker responses tended to consider AO1 and AO2 only with little or no AO3 content and generally gave far more vague statements with inaccuracies throughout.

Evaluate the extent to which human behaviour, such as the ability to drive, can be explained by biological psychology. ~~brain injury~~ ~~pseudo~~ psychodynamic hormones

You must make reference to the context in your answer.

(12)

Biological ~~ps~~psychology suggests that hormones play a role in aggression. Testosterone is a male ~~and~~ androgen that ~~ssss~~ is involved in the development of the limbic system, including the amygdala and hypothalamus as well as the fight or flight response. As males have more testosterone, especially during adolescence such as 17 years old, they may display more aggression. This would explain why Oscar and Jay failed their theory and practical driving tests and Nishka didn't as they have higher testosterone levels and therefore may be more reckless during driving tests and they ~~te~~ are more prone to aggression. Furthermore, Nishka may have higher cortisol

levels than Oscar and lay which is a hormone that's involves in managing the stress response - Therefore, she can may be calmer under pressure and less erratic, making her more able to drive safely.

Hormone theory can be said to be scientific. It uses objective measures such as ~~chromosome~~ blood tests which can be replicated if required. Therefore, this explanation has high scientific credibility due to its use of valid and reliable measures.
That said, it's a reductionist explanation. It ignores the role of individual differences in the development of aggression such as emotion and upbringing. Therefore, it's an over-simplified explanation and doesn't look at the 'bigger picture'.

Hormone theory takes into account gender differences in aggression. It suggests an explanation for the higher crime rate and higher levels of aggression in males as they have more testosterone. Therefore, it can be ~~unscientific~~ beneficial when applied to real life.

that said, hormone theory may be socially sensitive. It could be seen as assuming all males are prone to aggression, putting a negative label on them, and suggests that biological treatment could be put in place that ~~xxx~~ brings up issues in social control. Therefore, researchers must be responsible when conducting experiments involving this theory as they may have wider negative implications.

An alternate biological explanation is the psychodynamic approach. This would suggest that Jay may have failed his driving test due to an underdeveloped ego therefore, he is acting on his id and giving into the pleasure principle, gaining immediate gratification for his urges, suggesting this is why he speeded and failed his first attempt.

A non-biological explanation may be social learning theory. As Jay engays watching formula 1, he may see his role models speeding and gaining rewards for it - therefore, he may be more motivated to speed himself

so he can imitate them and gain these same rewards.

In conclusion, the hormone theory may help explain Jay's behaviour as due to high testosterone levels. However, hormone theory is reductionist and socially sensitive - therefore, the psychodynamic approach or social learning theory may better explain Jay's behaviour.

ResultsPlus
Examiner Comments

This response was awarded Level 2 – 5 marks.

The AO1 was judged as level 2 – Demonstrates mostly accurate knowledge and understanding.

The AO2 was judged as level 2 – Line(s) of argument occasionally supported through the application of relevant evidence from the context (scientific ideas, processes, techniques & procedures).

The AO3 was judged as level 2 – Candidate produces statements with some development in the form of mostly accurate and relevant factual material, leading to a superficial conclusion being made.

Question 6

Question 6 was an extended open response question with the 'Assess' taxonomy which targets both AO1 and AO3 content. AO1 was looking for knowledge and understanding of socially-sensitive research, including psychological theories, studies, methods and AO3 was analysis, interpretation, and evaluation of socially-sensitive research using content from psychology and the implications of this, leading to judgements/conclusions of how far they could be considered socially sensitive.

Assessment of this question was through a levels-based mark scheme where a 'best-fit' approach was used; deciding which level most closely describes the quality of the answer. Each AO was judged separately and where the components met the requirement for the level fully (and perhaps has elements of the level above), then marks were awarded at the top of the level. Where the components met the level but only just, they were awarded marks at the bottom end of the level. When a response was imbalanced (i.e. one AO was stronger than the other) a compromise was found. Consideration was also given regarding this question requiring greater AO3 content than AO1 (8/20 to AO1, 12/20 to AO3).

Candidate performance on this question was higher than in previous series in spite of the debate being assessed, which some may consider more difficult than other debates on the specification. This may have been due to the advance information provided. It was noticeable that candidates wrote a lot of content in general for this question, with varied quality. Candidates were clearly more prepared than usual but with varying degrees of success. The best responses gave an imbalanced response, with a greater focus on AO3 content than AO1, and explored their arguments with greater depth and complexity. They often considered the value of the research beyond its face value and explored the implications far more. Weaker responses generally gave superficial statements with arguments that had no depth beyond simplistic points regarding how socially sensitive the studies may be considered.

6 Assess the impact of socially-sensitive research in psychology.

① Rosenhan – ② Spitzer
③ Bradbury + Williams – BLM ④ [illegible]
⑤ Loftus + Palmer – Juries – Weinstein ⑥
↳ Flowers

(20)

There are ~~[illegible]~~ many studies in psychology which can be described as socially-sensitive, however it is often a question of whether their useful applications can justify the research.

One famous example is Rosenhan (1973) which looked at the reliability (~~[illegible]~~ consistency) and validity (how true it is) of diagnosis. Rosenhan got 8 confederates to ~~[illegible]~~ claim symptoms of voices saying "thud" and "empty" to 12 mental health institutions on the east and west coast of the US. Almost all confederates were admitted with the same diagnosis of schizophrenia, however despite the confederates being healthy and exhibiting no symptoms while in the institutions, it took an average of 19 days for them to be released, and up to 54 days for one individual. In the second part of the experiment, institutions ~~asked~~ ~~[illegible]~~ claimed that

Rosenhan had sent confederates and pseudopatients told some patients that they were certain were confederates. However, all of the patients were not sent by Rosenhan, as he had not sent any pseudopatients to any mental health institutions.

These two experiments demolished the trust the american public had in mental health institutions and led to decades of patients avoiding psychiatrists in fear of diagnosis. This led to fear to grow in other countries, and the same effect was also likely seen in the UK.

However, it could be argued that the useful applications outweigh the loss of trust. Spitzer et al (2015) showed 74 participants a case vignette which was the same as the pseudopatients in Rosenhan (1973), (Anyone who had heard of the study was excluded). Only 3 out of 74 psychiatrists were happy to give a diagnosis, with most citing a lack of enough information to give a valid

diagnosis. This shows that Rosenhan likely had a huge improvement on the validity of research, particularly on the diagnostic manuals used in the US such as the DSM, which is updated every 10 years.

Another example, this time in criminal psychology, is Bradbury and Williams. They investigated how racial makeup of the jury can affect conviction rates of black defendants by using a sample of real trials, and measuring the racial makeup of the jury, as well as the length of deliberation, the strength of the prosecution, they claimed that juries with a high number of white and/or hispanic jurors were more likely to convict black defendants.

This suggests that jury decision making is not fair or just, and that even race can affect how likely a defendant is to be convicted. This is extremely considering socially-sensitive, considering the fairly recent BLM riots, which focused

on the injustice shown towards black people by police officers. If it is shown that even in a ~~fair~~ trial they are at ~~the~~ a disadvantage, this can lead to even greater civil unrest.

However, the main difference between Bradbury and Williams and Rosenthal is that the exact applications are limited.

~~While we may be able to abolish juries~~

While we may be able to abolish juries in the future, such large scale changes take time, and people will go to trial knowing that they are at a disadvantage due to their race. In fact, juries may never be abolished, meaning that the main application of the study is not carried out.

Similarly, Loftus and Palmer looked at the reliability of eye-witnesses, and investigated leading questions (questions that provide an eye witness with information). When participants were asked how fast were two cars going when they – each

other in a video they watched, the mean estimate changed when the word used in the blank changed. "Smashed" gave a 10 mph higher speed than "contacted". Furthermore, people that saw the word "smashed" were more than twice as likely to report seeing broken glass in the video, even though there was none, when compared to those in the hit condition (7 to 16).

Loftus particularly shaped the socially-sensitive of her research, by acting as an expert witness in the Harvey Weinstein case. She told eye-witnesses that they were misremembering and their testimonies could not be trusted. Eye-witnesses would also have to see perpetrators go free that would otherwise go to jail not only dangerous for society, but traumatising for witnesses.

In conclusion, while Rosenhan (1973) shows that sometimes socially-sensitive research is necessary, Loftus and Palmer and Badkey and Williams show what happens when not enough is done to consider the implications of research. **(Total for Question 6 = 20 marks)**

ResultsPlus
Examiner Comments

This response was awarded Level 3 – 11 marks.

The AO1 was judged as level 3 – Demonstrates accurate and thorough knowledge and understanding.

The AO3 was judged as level 3 – Displays a logical assessment, containing logical chains of reasoning throughout which consider a range of factors. Demonstrates an understanding of competing arguments/factors but does not fully consider the significance of each which in turn leads to an imbalanced judgement being presented.

Paper Summary

Based on their performance on this paper, candidates should:

- Ensure they fully contextualise their responses when they are given a novel scenario and avoid generic statements throughout the paper.
- Read the questions carefully and include the necessary information as specified.
- Ensure they are using the required skills, such as fully justifying strengths, weaknesses, and improvements.
- Ensure 12 mark questions with a scenario have enough of each of the AOs and balance the amount of content given for each of the AOs.
- Make sure they give an imbalance on 16 and 20 mark questions with greater AO3 material than AO1.

Grade boundaries

Grade boundaries for this, and all other papers, can be found on the website on this link:

https://qualifications.pearson.com/en/support/support-topics/results-certification/grade-boundaries.html

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London WC2R 0RL.