

Examiners' Report

June 2025

GCE Psychology 9PS0 03

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 <https://www.edexcel.com> or <https://www.btec.co.uk>.

Alternatively, you can get in touch with us using the details on our [contact us page](#).



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 <https://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: <https://www.pearson.com/uk>.

June 2025

Publications Code 9PS0_03_2506_ER

All the material in this publication is copyright

© Pearson Education Ltd 2025

Introduction

The 2025 examination was the fourth full examination series since the pandemic. The performance of the candidates in this examination is summarised below, with advice on how to improve in future series.

Similar to previous series, the best responses for section A gave contextualised responses that were all specific to the novel scenario provided in the question. Where a question asked for 'one' strength/weakness/improvement for two marks, candidates provided full justification of their relevant identified point in relation to the study.

Once again, weaker responses were not fully-justified or were purely generic, so this remains a focus for centres to support candidates in future series. Additional focus to help support candidates would be with reinforcing an accurate understanding of the peer review process, what a review of studies entails, and the strengths of correlational research, because these were clear weaknesses for candidates in this series.

For section B, similar to previous series, the best responses focussed on considering relevant and appropriate points related to the unseen study Question (Q) 03 and on the required imbalance for the essay Q04. Higher-performing candidates had points that were accurate and in the context of the question, focussed on producing relevant and appropriate material, and showing clear understanding of what they were asked in the questions. Centres may again want to support candidates on the use of research evidence to support or refute application to the unseen study (Q03b), although this was noticeably better in the 2025 examination, which was pleasing.

Centres may also want to continue to help with reinforcing the imbalance towards Assessment Objective (AO3) for the essay (Q04). They should stress that there is only a limited amount of AO1 that can be credited, so an appropriate allocation of time needs to be offered to the AO1 and AO3 elements of the essay to achieve the most marks possible in the time allowed.

Regarding section C, the best responses once again considered the balance (for Q05), or imbalance (Q06) required for the command words and question types used. Knowledge and understanding (AO1) was relevant, accurate and thorough with varied ideas used as appropriate. Evaluation, analysis and assessment (AO3) was well-developed and led to appropriate judgements and conclusions throughout their response.

Centres may want to ensure candidates are using appropriate, accurate information regarding the questions asked and are not merely learning rote-learned essays in preparation for what they think a question may be. Centres may also want to continue to reinforce time allocations so that candidates continue to spend requisite time on all questions in the examination, giving them the best opportunity to gain the most credit for their work.

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 candidates for future 9PS0/03 examinations.

Question 1(a)

Candidates were required to explain a conclusion they could make using the data in the graph in relation to the scenario for Q01(a).

Overall, performance was split on this question. The best responses gave a clear, accurate conclusion and supported this using data from the graph. Weaker responses typically only wrote about data from the graph, with no conclusion identified.

There were 23 studies which considered the quantity of screen use and language skills. The results concerning average screen time (hours per week) from these studies are shown in **Figure 1**.

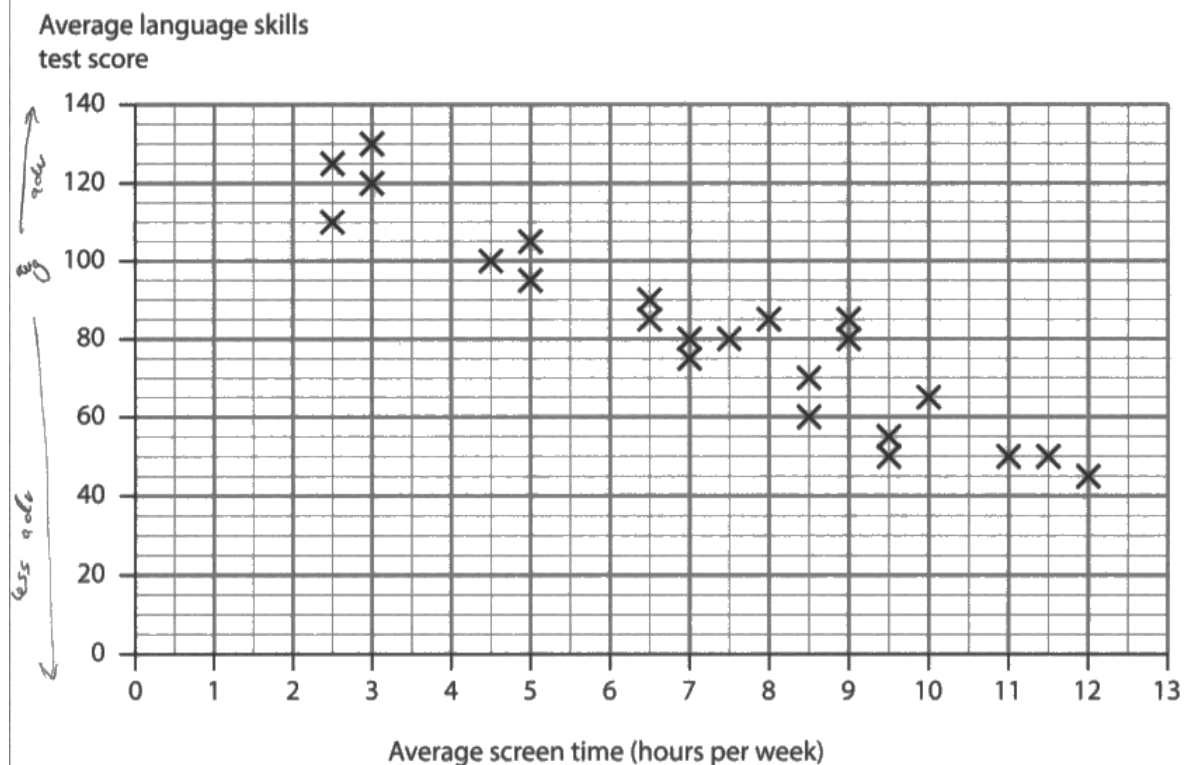


Figure 1

A scattergraph to show the relationship between average screen time (hours per week) and average language skills test scores

- (a) Explain **one** conclusion that could be made from the data in **Figure 1** in terms of the quantity of screen use.

(2)

As increased screen time leads to reduced understanding of language in children. This is shown because there is the average language skills test score decreases from 125 to 45 as average screen time per week increases from 2.5 hrs to 12 hrs.



This response is given 2 marks for identification of a conclusion and justification of the conclusion.

Question 1(b)

Q01(b) required candidates to calculate the percentage of studies, to one decimal place.

The majority gave the correct answer. The number of answers that rounded incorrectly or could not calculate the percentage accurately, despite having identified the correct number of studies, was surprising.

- (b) Calculate the percentage of studies that found children's language skills to be **greater** than the mean of 100, out of all the studies which considered quantity of screen use.

You must give your answer to **one** decimal place.

(1)

SPACE FOR CALCULATIONS

$$\frac{5}{23} \times 100 = 21.7$$

Percentage 21.7%



This response is given 1 mark for the correct percentage to one decimal place.

Question 1(c)

Candidates were required to explain a conclusion they could make using the data in the table in relation to the scenario for Q01(c).

Overall, performance was split on this question. The best responses gave a clear, accurate conclusion and supported this using data from the table.

Weaker responses typically only wrote about data from the table, with no conclusion identified. Repeating data from the table with no clear conclusion is something that should be avoided in future series.

There were 10 studies which considered the quality of screen use and language skills. The results concerning time spent viewing educational content are shown in **Table 1**.

Study	Average time spent viewing educational content (hours per week)	Average language skills score
A	1.5	121
B	0.1	49
C	2.2	138
D	0.3	64
E	0.2	55
F	1.6	112
G	0.7	88
H	0.5	76
I	1.1	102
J	0.9	98

Table 1

(c) Explain **one** conclusion that could be made from the data in **Table 1** in terms of the quality of screen use.

(2)

Children who spent more time viewing educational content were more likely to ~~have a higher~~ have better language skills. Shown by a higher average language skills score (121) in study A which had an average viewing educational content time of 1.5, compared to a lower ^{language skills} score of 55 in study E which had average 0.2 hours viewing educational content.



This response is given 2 marks for identification of a conclusion and justification of the conclusion.

Question 1(d)

For Q01(d), candidates were required to explain a conclusion they could make using the data in the graph in relation to the scenario.

Overall, performance was split on this question. Similarly to Q01(c), the best responses gave a clear, accurate conclusion and supported this using data from the table.

Weaker responses typically only wrote about data from the graph with no conclusion identified. As with Q01(c), repeating data from the table with no clear conclusion is something that should be avoided in future series.

There were 6 studies which considered the onset of screen use and language skills. The results concerning the age the children first began viewing screens are shown in **Table 2**.

Study	Average onset of screen use (in months)	Average language skills score
K	12	130
L	2	55
M	14	128
N	5	60
O	8	98
P	7	90

Table 2

(d) Explain **one** conclusion that could be made from the data in **Table 2** in terms of the onset of screen use.

(2)

The younger the child was when they started viewing screens, the lower their language skills were. This is shown as viewing screens at 2 months old had the lowest language skills score of 55 but having onset of screen use being 12 months, had a higher score of 130 which shows more advanced language skills.



This response is given 2 marks for identification of a conclusion and justification of the conclusion.

Question 1(e)

Candidates were required to explain one strength and one weakness of using studies that had been through peer review in relation to the scenario for Q01(e).

The best responses identified a strength and weakness and then justified each fully in the context of the scenario. Weaker responses only identified a strength and weakness, or gave only a relevant strength or weakness, in relation to the scenario.

The most common strength was the reduction of errors in the final review. The most common weakness was regarding the peer being biased. It was notable that a large proportion of candidates did not show an accurate understanding of the peer review process, so this could be a focus of support for centres for future series.

- (e) All of the research studies considered by the researchers had been through the process of peer review.

Explain **one** strength and **one** weakness of using studies that had been through the peer review process in the screen use and language skills study.

(4)

Strength

Peer reviewing a study allows it to be checked via the concerned validity of multiple researchers coming to the same conclusion, such as all agreeing that prolonged screen time negatively affects language development. This is a strength as it allows studies to be thoroughly checked, as the studies are reliable and valid, which is a strength as it means the studies into the onset of screen time or onset of screens are more valid.

Weakness

One weakness is that the peer review will deal with secondary data as they have not usually only assess studies they believe to be appropriate or beneficial to the conclusion proposed, such as that prolonged screen time affects language skills negatively. This is a weakness because it leads to publication bias, where studies that may reach the intended conclusion or are not published in general or studies that agree with the study into how screen time affects age of onset of screen time negatively affects language skills and therefore also makes the studies less valid.



This response is given 4 marks for identification of the strength and weakness and for the justification of each.

Question 1(f)

For Q01(f), candidates were required to identify an appropriate improvement for the study and then justify this fully in context.

The best responses offered a relevant and appropriate improvement and then justified it in the context of the study. There was a wide range of improvements offered by candidates.

Weaker responses gave an inappropriate improvement, focussed on a weakness of the study only, gave a generic response, or where an appropriate improvement was suggested, there was no relevant justification to support their improvement.

Similar to Q01(e), it was notable regarding the number of responses that did not show an accurate understanding of a review of studies, despite one being compulsory on the specification, which limited credit that could be awarded.

(f) Explain **one** improvement that could have been made to the screen use and language skills study.

(2)

one improvement that could be made to the screen use and language skills study is that researchers^{or} could use more than one ~~assessment~~ test or measurement for the children's language skills, such as using secondary data by assessing ~~the~~ the children's school english ~~or~~ test scores. This is because by only using ~~the~~ the Picture Vocabulary Test to measure language skills the ~~researcher~~ study may lack validity so it is better to gather more data to have a more accurate and detailed portrayal of language ^{skills}.



This response is given 2 marks for identification of an improvement and justification of the improvement.

Question 2(a)

Q02(a) required candidates to calculate the mean score to one decimal place. The vast majority gave the correct answer.

The results of the mobile phone use and well-being study are shown in **Table 3**.

Participant	Mobile phone use score (0–44)	Satisfaction with life score (5–35)	Well-being score (0–25)
A	42	10	2
B	38	12	5
C	26	17	9
D	8	23	22
E	16	19	18
F	12	22	22
G	16	15	15
H	33	9	6

Table 3

- (a) Calculate the mean score for the participants' satisfaction with life using the data in **Table 3**. You must give your answer to **one** decimal place.

(1)

SPACE FOR CALCULATIONS

$$\begin{aligned}10 + 12 + 17 + 23 + 19 + 22 + 15 + 9 &= 127 \\127 \div 8 &= 15.875 \\&= 15.9\end{aligned}$$

Mean 15.9



This response is given 1 mark for correct calculation of the mean to one decimal place.

Question 2(b)

For Q02(b), candidates were required to complete the table and calculate the Spearman's correlation coefficient to three decimal places.

The best responses completed the table correctly and indicated their coefficient clearly. Weaker responses only completed part of the table correctly, typically only the first and/or second difference columns or did not complete any part of the table correctly.

A relatively common error was for candidates to minus 1, not understanding this meant the final coefficient went to a negative value, which limited them to 3 out of the 4 available marks.

- (b) The researchers decided to use a Spearman's rho test to see if there was a relationship between mobile phone use and well-being.

Complete **Table 4** and calculate Spearman's rho for the relationship between mobile phone use and well-being.

You must give your answer to **three** decimal places.

(4)

Mobile phone use score (0-44)	Rank	Well-being score (0-25)	Rank	d	d ²
42	8	2	1	7	49
38	7	5	2	5	25
26	5	9	4	1	1
8	1	22	7.5	-6.5	42.25
16	3.5	18	6	-2.5	6.25
12	2	22	7.5	-5.5	30.25
16	3.5	15	5	-1.5	2.25
33	6	6	3	3	9
$n = 8$				Total:	165

Table 4

SPACE FOR CALCULATIONS

$$\begin{aligned}
 1 - \frac{6 \sum d^2}{n(n^2-1)} &= 1 - \frac{6 \times 165}{8(8^2-1)} \\
 &= 1 - \frac{990}{8 \times 63} \\
 &= 1 - \frac{990}{504} \\
 &= 1 - 1.964 \\
 &= -0.964
 \end{aligned}$$

Spearman's rank correlation coefficient -0.964



This response is given 4 marks for correct calculation of the Spearman's correlation coefficient to three decimal places.

Question 2(c)

Q02(c) required candidates to explain a weakness of using primary data in the context of the study.

The best responses identified a weakness in context and then justified this fully in relation to the scenario. Weaker responses tended to give a generic response or only identified a weakness in context. The most common weakness was that it would be more time-consuming than using secondary data.

(c) The researchers in the mobile phone use and well-being study used primary data.

Explain **one** weakness of using primary data for the mobile phone use and well-being study.

(2)

One weakness of using primary data ~~is that~~ for the mobile phone use and well-being study is that it is more time consuming than gathering secondary data. Researchers have had to gather participants from the German University and get them to complete scales regarding their mobile phone use and well-being. As opposed to using existing research. As a result, it is likely to take a significant amount of time to collect data for the mobile phone use study, delaying across the results and conclusion, proving a weakness for using primary data.



This response is given 2 marks for identification of a weakness and justification of the weakness.

Question 2(d)

Candidates were required to explain one strength and one weakness of using correlational research in relation to the scenario for Q02(d).

The best responses identified a strength and weakness and then justified each fully in the context of the scenario. Weaker responses most commonly showed only accurate understanding of a weakness or gave only an identified strength or weakness in relation to the scenario.

The most typical strength was either that they are more practical/ethical or they can be preliminary research.

The most common weakness was a lack of cause-and-effect. Generic responses were less common this series, which was positive, and centres may wish to continue supporting candidates with having contextualised responses.

(d) Explain **one** strength and **one** weakness of using correlational research for the mobile phone use and well-being study.

(4)

Strength

one strength is that correlational research can be useful to gain data in ways that would otherwise be impractical and unethical to collect. For example, it would be unethical to force someone ~~they~~ use their mobile phone for more than 8 hours a day to simply see how it will affect their well-being as it could cause harm to their health. This means correlational research is useful to find information on mobile phone use and well-being in an ethical and psychologically practical way.

Weakness

one weakness of correlational research is that no ^(causal) cause-effect relationship can be established. This is because the research cannot indicate whether a high mobile phone use will impact well-being or whether a good level of well-being will impact mobile phone use. This means the conclusions made from the study on phone use and wellbeing are incomplete as it is unclear which variable affects the other.



This response is given 4 marks for identification of the strength and weakness and for the justification of each.

Question 3(a)

Q03(a) required candidates to explain a practical application of the results of the study.

The best responses identified a relevant application in context and then justified this fully in relation to the data from the table. Weaker responses tended to give a conclusion and not a practical application or only identified a practical application in context.

The most common practical application was for teachers to praise children during PE lessons in school.

The results of the preschool children and PE study are shown in **Table 5**.

		Average number of steps			
		More active male teacher with praise	More active female teacher with praise	Less active male teacher with no praise	Less active female teacher with no praise
Overall		1560	1555	765	760
Gender	Boys	1810	1600	1010	810
	Girls	1310	1510	520	710

Table 5

(a) Explain **one** practical application of the preschool children and PE study, using the data from **Table 5**.

(2)

One practical application of the preschool children and PE study would be that PE teachers should be more active and give praise to encourage preschool children to be more active and take more steps. This is shown as children with more ~~encouraging~~ ~~more active teachers~~ active teachers did more steps with 1560 and 1555 overall compared to 765 and 760 which is lower for less active teachers and no praise.



This response is given 2 marks for identification of a practical application and for justification of the practical application.

Question 3(b)

For Q03(b), candidates were required to apply social learning theory to the findings of the study and then use research evidence to support or refute how social learning theory could account for the findings of the study. Typically, candidates wrote more than was necessary, with good knowledge of social learning theory shown, but sometimes they did not link this to the findings of the study.

The general performance on this question was much better than in previous years, which was pleasing, and some candidates showed some very good skills in relation to the requirements of the question. However, some responses made no relation to the findings of the study or to any research evidence, thereby not answering any aspect of the question asked.

The best responses used the specific findings from the scenario and clearly and appropriately linked content from social learning theory to account for the findings. These responses tended to use specific data from the results table. They then used appropriate research evidence to justify their points. The Bandura studies were the most commonly-used for supporting research evidence, with studies related to biological psychology as an opposing argument also fairly frequent.

Weaker responses focussed on application only and tended to write with much knowledge and understanding about the theory and then apply this perhaps once to the study's findings, in a more general manner. Some responses did not link social learning theory to the findings of the study or when considering AO3 gave alternative theories, not research evidence, as specified in the question.

(b) Using research evidence, explain how far social learning theory can account for the findings of the preschool children and PE study.

SLT

(6)

One way social learning theory can account for the findings of the preschool children and PE study is through observation and imitation. For example, in this case students observed their PE teacher ^(role model) walking and getting their steps in and so, the children learned this behaviour and followed the teacher by imitating this behaviour, which could explain why those teacher who had a higher average number of steps had children performing better and also having ~~the~~ number of average steps e.g. boys - 1810 and girls - 1310 (more active male teacher with praise). ~~whereas~~ This is supported by Bandura's study, where the study found that children learned and imitated the aggressive behaviour when showed by the live role models acting aggressive towards the bobo doll and so, could ~~get~~ account for this findings as well. This is a strength as social learning theory can account for the findings of the study.

Another way social learning theory can ~~be~~ account for the findings of this study is through same-sex affect and vicarious reinforcement. For example, ~~if~~ when children were praised, they were more motivated to repeat the action and do more steps as when they were not praised, this can be seen in the data as well e.g. less active male teacher with no praise had a overall score of 765 average number of steps vs when praised - 1560 ~~the~~ average steps (praise + ^{active} male teacher). Also, children

are more likely to learn a behaviour when they are similar to each other
e.g. boys had a average number of steps - 1810 when
it was a male teacher praising & active vs 1600 when a
active female teacher was praising. This is supported by Bandura
when male models were acting aggressive, ~~the~~ boys were more
~~likely to be~~ ^{also aggressive} towards the bobo doll. This is a strength a SLT can
explain the findings of this study.

However, an alternative explanation could be that boys
have a higher testosterone levels than girls and so, this
could be why they had a better average number of
steps than girls did e.g. boys - 1810 vs girls - 1310
and this wasn't the only case for when instructed by ^{active} male
/ female teacher but, both and so,

(Total for Question 3 = 8 marks)

↳ it could be more of a biological explanation rather
than learning theories. This is a weakness as
it cannot explain this
factor into
this study.



This response is given 5 marks for:

- AO2 application of observation and imitation of role models to the findings (1)
- AO3 Bandura as evidence to support their AO2 about teachers as role models (1)
- AO2 application of external motivation to the findings (1)
- AO2 application of same-sex imitation to the findings (1)
- AO3 Bandura as evidence to support their (1)
- AO2 same-sex role model(1)

No credit for biological psychology as an alternative explanation because it is not research evidence as specified in the question.

Question 4

Q04 was an extended open-response question with the 'Evaluate' taxonomy that targets both AO1 and AO3 content.

AO1 was looking for knowledge and understanding of the studies or ethics. 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 described the quality of the answer.

Each AO was judged separately and where the components met the requirement for the level fully (and perhaps had elements of the level above), then marks were given at the top of the level. Where the components met the level but only just, they were given marks at the bottom end of the level. When a response was imbalanced (ie 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).

Candidate performance for this question maintained a similar standard to previous series, albeit with a slightly higher performance than the equivalent question in 2024. The topic of ethics may account for this because some candidates showed competent understanding of the ethical guidelines. In general, good knowledge of Rosenhan and Raine was evidenced by candidates.

The best responses gave accurate and thorough knowledge and understanding, usually of the studies, and then considered the ethical nature of the studies accurately, in a methodical way. Typical ethical issues cited by candidates included informed consent, deception and protection of participants.

The very best responses weighted their response far more towards the analysis and evaluation (AO3) and made detailed, accurate judgements and conclusions regarding how ethical each study could be considered.

Weaker responses focussed the majority of their response on knowledge and understanding (AO1) of the studies with few evaluative or analytical statements (AO3) and showed limited understanding of the ethical issues.

Some responses still focussed erroneously on the pseudopatients as the participants, which should be avoided in future series.

Similar to 2024, it was again noticeable that some candidates focussed so much on the AO1 that they achieved beyond Level 4 for this element but may have only reached the bottom of Level 2 for their AO3. Consequently, their overall mark was limited by the imbalance in their response towards their AO1 instead of their AO3.

These candidates would have been better to reduce the AO1 they offered (they could have still gained Level 4 for this element) and focussed more on increasing the amount and detail of their AO3 content, to achieve a stronger overall response.

The weakest responses gave very generic and vague information regarding the studies with little to no evaluation.

9:40 → 10:00

4 Evaluate Rosenhan (1973) and Raine et al. (1997) in terms of ethical issues.

(16)

The BPS ethical guidelines (1990) outline 4 main principles, integrity, respect, responsibility and competence. These guidelines must be followed for a study to be deemed ethical, and if not followed may put participant (ppts) in unnecessary harm.

In Rosenhan's (1973) study, 8 pseudopatients gave fake symptoms (auditory hallucinations) of the words empty, hollow, and flat to be admitted into 12 different psychiatric hospitals across 12 states, in order to test whether psychiatrists could reliably tell the sane from the insane. They found that on average of 8-52 days, 7 were given an SZ diagnosis, 1 bipolar and that no staff could tell they were fake patients, concluding that psychiatrists could not tell the sane from the insane.

On one hand, Rosenhan's study could be considered ethical as all pseudopatients had to give fake names & employment status upon admission to the psychiatric hospital. This was in order to protect the pseudopatients' identity when looking for future job prospects. This could be ethical as (although it hadn't been created yet)

follows the ethical guideline of responsibility, as it ensures that the participant in the study does not affect their future.

However, the procedure of Rutterman's study involves deceiving hospital staff across 12 different hospitals which breaches the integrity BPS ethical guideline.

As a result, this could cause distress for nursing staff at the hospitals as they may question their own competency in the career.

Despite this, deception was completely necessary for the aim of the study, as the aim was to test if staff could tell whether the past ^(MHOs) ~~present~~ patients and actually have mental health disorders or not, and also led to the useful application of an updated version of the DSM, which ^{correctly} ~~more~~ diagnosed patients, giving future diagnosis to be more accurate.

In Raine's (1997) study, he aimed to see if there was a correlation between cortical and sub-cortical brain activity and aggression, using 41 ~~not~~ not guilty due to reason of insanity (NGRI) murders or ~~not~~ non-murders and controls. There were matched for age, sex and history with 1/2 and had all been medicated for 2 weeks. Raine conducted PET scans to measure

brain activity after injecting ppts with an FDG tracer and conducting a control performance task (CPT) for 32 mins. The panel test results had lower activity in the corpus callosum and the amygdala but higher activity in the right hemisphere, concluding that biological predisposition to violence to exist but the environment trigger the behavior.

Laine's study may be considered ethical as he did not disclose the identity of any of the NCRIs he was studying, following the BPS guideline of protection. This allows each individual to not be stigmatized for their brain activity or behavior and the likelihood of the person being labelled as abnormal. Equally, by protecting their identity Laine does not attempt to justify the crime they committed with brain activity which could invalidate the victim's family's grief or suffering, causing them psychological harm. However Laine's study could be deemed unethical as he prevented each ppt from taking medication for two weeks prior to the PET scan, which may alter mood and behavior. This may make it difficult for ppts to manage any symptoms of MHPs which could cause unnecessary harm, breaching the guideline of responsibility.

While this was vital to monitor brain activity in a controlled manner, these films are still only representative of NCRI, and may have caused too much emotional distress to ppts to justify carrying out the study.

However, earlier films have been significantly useful for NCRI in court, e.g. Donita Page, who used the fact that his postal boxes were mislabeled to reduce his sentence from the death penalty to life in prison.

Therefore overall, psychological research such as Rutter (1973) and Raine (1997) have become more ethical since the introduction of the BPS ethical guidelines in 1990, allowing ppts to be protected and kept from unnecessary harm. However, due to the introduction of the guidelines, ~~past~~ advancements in psychological research are limited, as procedures deemed unethical, despite the advancements that they could promote, are not allowed to occur.



AO1: Level 3

Demonstrates accurate knowledge and understanding.

AO3: Level 3

Arguments developed using mostly coherent chains of reasoning leading to a conclusion being presented. Candidates will demonstrate an understanding of competing arguments but evaluation may be imbalanced.

Overall, the response is placed in Level 3.

Level 3

Total: 12 marks

Question 5

Q05 was an extended open-response question with the 'Evaluate' taxonomy, with a scenario that targets AO1, AO2 and AO3 content.

AO1 was looking for knowledge and understanding of social psychology, AO2 was application to the scenario given in the question about the scenario involving Kristie and Dominique. AO3 was analysis, interpretation, and evaluation of the ideas presented from social 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 to decide which level most closely described the quality of the answer.

Each AO was judged separately and, where the components met the requirement for the level fully (and perhaps had elements of the level above), marks were given at the top of the level. Where the components met the level but only just, they received marks at the bottom end of the level. When a response was imbalanced (e.g. 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 slightly better than in previous years. The best responses focussed on offering a balance of AO1, AO2 and AO3 content with knowledge and understanding of social psychology, application of this to the scenario, and then evaluation and analysis leading to judgements and conclusions.

The most common content used by candidates in their response included agency theory and social impact theory for their AO1 and AO2. Some also used social identity theory and realistic conflict theory. Learning theories were used frequently as part of the AO3, as were research studies for social psychology, such as Milgram, Hofling, Rank and Jacobson, Sedikides and Jackson, and Tajfel's research.

Weaker responses focussed very narrowly on one or two areas of social psychology using more generic, vague statements with little or no evaluation or analysis in their response. At times, they focussed largely on how other areas of psychology could explain the scenario instead of social psychology, losing the focus of their response.

There was commonly an imbalance towards AO2 content in particular, which largely recycled information from the scenario instead of using the knowledge and understanding to apply in a focussed and engaged way to the novel scenario presented.

- 5 Kristie is at school in her biology class, sitting next to her friend Dominique. Dominique regularly cheats on class tests and Kristie has witnessed this but always does her tests honestly. The teacher announces the class will have a progress test the next week.

Dominique is very popular and frequently tells other students what to do and they always follow what she says. Recently, she told Kristie to trip up another student as they were going down the stairs and she did so. As a result, Dominique invited Kristie to an exclusive party.

Dominique sees the biology teacher leave a copy of the forthcoming progress test on her desk as they leave the lesson. When they are alone, Dominique tells Kristie to sneak into the biology teacher's classroom during lunchtime and take photos of the forthcoming progress test. She tells Kristie that if she does, she will be rewarded like she was previously.

Kristie sneaks into the classroom and gets caught when taking photos of the test by the biology teacher. When asked why she has done this, Kristie says she was only following what another student told her to do.

Evaluate the extent to which human behaviour, such as the events at Kristie's school, can be explained by social psychology.

You must make reference to the context in your answer.

(12)

One theory is agency theory, which states that when an individual is faced with an order by a legitimate authority figure (LAF), they make the agentic shift from the autonomous state (where they have free will and responsibility for the choices of their actions) to the agentic state (where they blindly obey the LAF, acting as their 'agent', and all responsibility for the consequences of their actions lies with the LAF).

Kristie may see Dominique as a LAF due to her high status of being popular so that when Dominique gives the orders to trip up another student when they were going down the stairs, Kristie makes the agentic shift into the agentic state, ~~because~~ thus tripping the other student up.

One strength is supporting evidence from Milgram (1963) who found that 100% of pps went to 300V when asked to electrically shock a learner (unfed state) when they got a word pair task wrong and 65% went to the maximum 450V. Therefore, this supports the idea that Krutie will make the agentic shift into the agentic state when given an order by the LAF (Dominique) to trip another student down the stairs.

However, Milgram was a laboratory experiment that took place within a controlled environment and so the results into obedience may not be generalisable to the situation with Krutie at school, where she tripped another student up in the agentic state, thus limiting the ability to explain this behaviour through social psychology.

Another theory is social impact theory, which states that the higher strength, immediacy and number that the source has, the more obedience the target will show. Strength is the perceived power of the source, gained from uniform, age or status.

~~For~~ Immediacy refers to if the orders are given face to face (spatial) and must be carried out immediately (temporal).

Number refers to the amount of sources present. The mathematical equation $i = f(SIN)$ demonstrates how these three social forces interact.

Krutie will perceive Dominique as having high strength due to her status of being a friend and being popular. Dominique's orders to sneak into the biology teacher's classroom and take photos of the

progress left her given face to face and expected to be carried out that lunchtime, which also increased the likelihood that Kristie will obey her orders.

One strength is supporting evidence from Sedikides and Jackson, who found that 78% of ppts obeyed a zookeeper's orders to not lean against the fence when they were in uniform compared to the 37% who obeyed when the zookeeper was in ordinary clothes. Therefore, this supports the idea that Dominique being higher strength will lead to Kristie obeying her orders to take photos of the upcoming progress test.

However, Dominique wasn't in a uniform, which limits the generalisation of this evidence to this situation. Instead, an alternative explanation is operant conditioning, as Skinner (1978) found that rats and pigeons in a Skinner's box were more likely to press a lever when they received a food pellet each time they pressed it. Therefore, this supports the idea that Kristie may take photos of the test and put the thumbs down the chain as this receives rewards and positive reinforcement for it, such as being invited to an exclusive party, which challenges the idea that social psychology can fully explain her behaviour.

In conclusion, social psychology can explain the events at Kristie's school to some extent, as evidenced by supporting research. Competitions, for the limited reward of gaining an invite to the exclusive party may make Kristie display more behaviour, such as kicking students down the stairs in order to earn the reward, according

to realize conflict than. However, to consider social psychology alone is to ignore other factors, such as neurotransmitters, which Kristie may feel good about being a student dancer, traits and acts, rebellious, which leads to an increase of dopamine, which makes her feel a sense of reward and motivation to continue her behavior, thus social psychology can only explain to some extent.



AO1: Level 3

Demonstrates accurate knowledge and understanding.

AO2: Level 4

Line(s) of argument supported throughout by sustained application of relevant evidence from the context (scientific ideas, processes, techniques or procedures). Demonstrates the ability to integrate and synthesise relevant knowledge.

AO3: Level 4

Displays a well-developed and logical evaluation, containing logical chains of reasoning throughout. Demonstrates an awareness of competing arguments, presenting a balanced conclusion.

Overall, the response is placed in Level 4.

Level 4

Total: 10 marks

Question 6

Q06 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 how psychology has developed over time in relation to two of the four applications listed, including psychological theories, studies and methods.

AO3 was analysis, interpretation, and evaluation of how psychology has developed over time using content relevant to two of the four applications listed and the implications of this, leading to judgements/conclusions of how far they could be considered a means of progression.

Assessment of this question was through a levels-based mark scheme where a 'best-fit' approach was used to decide which level most closely described the quality of the answer.

Each AO was judged separately and, where the components met the requirement for the level fully (and perhaps had elements of the level above), marks were given at the top of the level. Where the components met the level but only just, they were given marks at the bottom end of the level. When a response was imbalanced (e.g. 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).

General performance on this question was very similar to previous years. As with previous years, candidates tended to write a large amount for this essay, usually recounting a large volume of what they had learned, sometimes with little focus on how psychology had developed over time.

The two most frequent applications used were clinical and criminological.

For clinical, candidates wrote most commonly about diagnosis in relation to the Diagnostic and Statistical Manual of Mental Disorder (DSM)/International Classification of Diseases (ICD) and treatments for mental health issues such as the advancement in drug treatments and use of non-biological treatments (such as Cognitive Behavioural Therapy (CBT)).

For criminological psychology, changes in eyewitness testimony in light of research and explanations for crime/anti-social behaviour were typical, with treatments for offenders also fairly frequent. A variety of content was used in relation to child and health psychology.

Some candidates used content not linked to the applications listed and looked as if they were only reproducing a rote-learned essay, with little consideration of the actual question asked, which should be avoided in future series.

The best responses considered the imbalance required towards the AO3 content and offered accurate and thorough knowledge and understanding, as well as a well-developed and logical assessment.

The AO1 presented tended to be a mix of accurate understanding of theories, studies and methods. The AO3 focussed mainly on how far knowledge could be considered to have progressed and the positive and negative implications of the changes over time in psychology. It

was notable how those candidates who achieved higher marks showed a higher degree of awareness both of how things had progressed over a timeline and the implications of any changes on various stakeholders.

Weaker responses tended to focus on recalling anything they could remember from the applications, rather than focus on how knowledge and understanding has developed over time.

Similar to previous years, whilst much was written by the majority of candidates, weaker responses tended to write unfocussed, inaccurate statements and a considerable amount of vague, generic content. Once again, those candidates who received the lower marks also tended to focus more on AO1 than AO3 content, rather than the imbalance towards AO3, which would have enabled a higher mark overall.

6 Assess how psychological understanding has developed over time, with reference to any **two** of the following applications to psychology:

- Clinical psychology $\begin{matrix} \swarrow \text{dopamine hypothesis} \\ \searrow \text{CBT} \end{matrix}$ $\rightarrow$ combined approach
- Criminological psychology $\rightarrow$ biological explanations $\times \times$
- Child psychology $\rightarrow$ ADT treatment
- Health psychology.

(20)

Psychological understanding has greatly developed over time in all psychological approaches but particularly clinical and criminological psychology. In clinical, the dopamine hypothesis has been renewed since the 60s 50s and 60s all the way up to 2009, with new information consistently ~~been~~ being added over time. Studies such as Renschen (1973) have contributed to the renewal of dopamine and the approach to treatment has greatly changed. In criminological, explanations of crime have been suggested and debated ($\times \times \times$) over time. This has led to improvements in treatment from both a biological and cognitive approach.

Rossm (1966) initially suggested that the dopamine hypothesis ~~a~~ explained schizophrenia in terms of ~~high~~ ^{high} dopamine levels caused by too many D2 receptors and low levels of beta

hydroxide. This led to the development of the first antipsychotics that blocked D2 receptors lowering dopamine activity in the synapse. This is significant as the early versions of the dopamine hypothesis paved the way for the further additions and provided a foundation for treatments.

Davis (1991) then suggested that lower levels of dopamine in some parts of the brain are caused by higher levels in places such as the striatum which leads to both positive and negative symptoms. Negative symptoms are linked to low levels of D2 receptors in the meso-limbic pathway whilst a high level of D2 receptors in the meso-cortical pathways leads to positive symptoms. This contribution to the hypothesis was highly significant as it led to the development of atypical drugs which were found to be effective for the 60% of Sz patients who found no relief from typical antipsychotics e.g. Chlorpromazine.

The most recent adaptation of the hypothesis is Flores and Kapur (2009) dopamine dysregulation that argues variation in dopamine levels in presynaptic neurons can trigger psychosis instead of looking purely at the receptors.

Whilst this is still a significant addition to the dopamine hypothesis, it has not led to any ground breaking developments in medication. So may not be the most impactful of the 3 major changes.

For a while in clinical, medication was the preferred treatment for SZ alongside securing the most extreme cases. Festerh (1973) helped to alter this method of treatment and diagnosis by highlighting that doctors were too quick to assume sickness over health and see any behaviour as a symptom. Whilst now it can be argued the opposite is true, many SZ patients receive medication from doctors as they often assume health over sickness which is a negative impact of this psychological knowledge changing our time.

Psychologists tend to take a more holistic treatment approach for those with SZ (and OCD) as a combined approach of CBT and medication has been proven to be most effective at reducing symptoms at $p < 0.001$ (POTS 2004). This is extremely important as this development ensures patients are getting the best care they need.

Explanations of crime have seen significant change over time. A prime example is XYY. Initially, researchers believed that XYY led to ~~was~~ a predisposition to crime due to the assumption ~~that~~ that an extra Y chromosome increases testosterone and therefore aggression. Jacobs (1965) found that the prevalence of XYY in the general population was 0.2% compared to 3% in the prison population leading to the conclusion XYY increases the likelihood of crime. This conclusion however was debunked by a following study that found no significant difference in the rates of XYY in prisons vs the wider population and instead suggested the results were due to the fact XYY decreases IQ and makes the person physically bigger so they are more likely to be caught. This progression in explaining crime is significant as it prevented XYY people from being labelled as criminals ~~and~~ when it simply was not true. Additionally it shows psychology's progression as a science and the adaptations criminologists' psychology has made when new information has been presented.

By adopting the biological explanation's newer treatments have been brought about to help offenders with testosterone. Testosterone has been a hormone identified as ~~an offender~~ in high levels for offenders suffering with paraphilia who have committed sexual crimes. Androgen deprivation treatment (ADT) blocks the testosterone hormone which lowers the production of testosterone and breaks down existing testosterone in the bloodstream. A case study, Michael Ross, was a serial killer given Depo-provera (ADT) and was found to have reduced his testosterone levels by 20x. Continued research on biological treatments for crime have recently suggested SSRIs may be helpful in reducing the compulsive symptoms of paraphilia and provide a more cognitive approach paired with CBT. This is significant as the progression in treatment for offenders greatly benefits society.

Overall, clinical psychology has greatly changed over time. The dopamine hypothesis has been revisited and

focus leading to development in treatment.

↳ Community Psychology is a slightly newer approach but still has had significant attention in its understanding of what causes crime and how to treat it.



AO1: Levels 4 and 5

Demonstrates accurate and thorough knowledge and understanding.

AO3: Level 4

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.

Overall, the response is placed in Level 4.

Level 4

Total: 14 marks

Paper Summary

Based on their performance on this paper, candidates should:

- Ensure they understand fully what a peer review process and review of studies involve, as well as the strengths of correlational research
- Accurately complete the table for Spearman's and be able to calculate the final coefficient correctly
- Avoid generic statements when given a novel scenario. All responses should be in the context of the study given
- Justify fully any points given if the question is offering two available marks, such as for a strength, weakness or an improvement
- Balance AO1, AO2, AO3 and include the requisite amount of evaluation if given a 12-mark question with a scenario, such as Q05 on this 2025 paper
- Learn information appropriately and answer the questions asked, not relying on rote-learned essays and predicting what the questions will be
- Focus more on AO3 content on evaluate or assess questions worth 16 marks or 20 marks with no scenario, such as Q04 and Q06 on this 2025 paper, whilst allowing enough time fully to attempt all questions on the paper

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>

