RESEARCH METHODS PAST PAPER QUESTIONS

154 patients who suffered from panic attacks were asked by a psychologist to take part in a clinical trial to assess the effectiveness of cognitive behavioural therapy (CBT). In order to select suitable participants for the trial, each patient completed a questionnaire which measured the severity of their symptoms on a scale of 1 (mild) to 10 (severe). 60 patients who had moderate symptoms with scores of 7 or 8 were selected to take part. They were randomly allocated to one of three conditions.

● Condition 1: Traditional cognitive behavioural therapy – this consisted of 12 one-hour sessions over a 12-week period.

● Condition 2: Brief cognitive behavioural therapy – this consisted of 5 one-hour sessions over a 5-week period with homework assessments.

● Condition 3: Control condition – patients were placed on a 12-week waiting list for traditional cognitive behavioural therapy. This group did not receive any form of treatment during the 12 weeks.

One week after finishing the course of therapy, patients in condition 1 and condition 2 completed the questionnaire for a second time. Patients in condition 3 completed the questionnaire for a second time at the end of the 12-week period. The three scores from this questionnaire were compared to see if there were differences in the severity of symptoms.

1. **Identify one variable which does not appear to have been controlled in this trial. Explain how this may have influenced the outcome. [3 marks]**

The length of time before second assessment of symptoms was 12 weeks in conditions 1 & 3 compared to 5 weeks in condition / Use of different therapists in conditions 1 & 2 / The interaction between sex of therapist and patient / Whether or not patients carried out homework tasks / Individual differences such as age and gender / Whether patients were receiving other forms of therapy or medication.

Eg: 'An extraneous variable was the length of time before second assessment of symptoms. This was 12 weeks in conditions 1 & 3 compared to 5 weeks in condition 2. This could have influenced the outcome of the study as the benefits of brief CBT have less time to show than the other two conditions. This could lead to its effectiveness being underestimated.

2. **What is meant by 'validity'? How could the psychologist have assessed the validity of the questionnaire used to measure the severity of symptoms? [4 marks]**

It could be assessed by: Taking another measure of symptoms from the same participants (eg Doctor or family member) and comparing the two sets of scores. If the scores agree, the questionnaire has high validity(concurrent validity) / Ask an expert(s) in the field to assess the questions to see if they are an accurate measure of panic attacks (content validity) / Assess how closely the questions relate to underlying theoretical constructs (i.e. how well they relate to panic symptoms) – construct validity / Less rigorous methods include looking at the questions to see if they appear valid 'on the face of it' (face validity).

Eg: 'Validity refers to whether or not the questionnaire measures what it is supposed to measure (1 mark). Concurrent validity would involve getting a family member to assess the symptoms (1 mark) and seeing how closely they match the score on the questionnaire (1 mark). If the two matched, the questionnaire would have high validity

3. **The psychologist asked the 60 patients for fully informed consent to take part in this trial. What should the psychologist have told the patients so that they were able to give their consent? [5 marks]**

In order to gain fully informed consent for this trial, patients should be informed of key information provided in the stem about the clinical trial. They will be allocated to one of the conditions and they may not receive therapy. If they do receive therapy it will be Cognitive Behavioural. The time period for the study (ie up to 12 weeks)

In addition, students could refer to other relevant ethical information such as: Data should be anonymised so they are not identifiable in the results / Patients should be made aware that they are free to withdraw themselves or their data from the clinical trial if need be / They may be asked to complete homework assignments outside the therapy sessions.

For five marks, students must cover the top three bullet points.

The psychologist used a statistical test to assess the differences in severity of symptoms between patients in Condition 1 and Condition 2. The difference between traditional CBT and brief CBT was found to be non-significant ($p > 0.05$).

Severity of symptom scores after therapy	Condition 1 Traditional CBT	Condition 2 Brief CBT	Condition 3 Waiting list
Mean	4.2	4.3	6.7
Standard deviation	1.2	2.4	1.3

4. **What do the data show about the effectiveness of the therapies for panic attacks? Refer to mean scores, standard deviations and the results of the statistical test in your answer. [6 marks]**

The mean scores show that both types of CBT lead to some improvement in symptoms and appear superior to being allocated to a waiting list • The SD score is larger for the brief CBT than the other groups, showing that the scores are more spread out in this condition than the other two groups. Hence, there was more variation in response to the short version of CBT • The statistical tests show no significant difference between short and long versions of CBT • The waiting list condition also demonstrated a minor improvement in symptoms. Students could refer to spontaneous remission here.

5. **In the discussion section of a report, researchers are expected to consider possible practical applications and implications of their research. Discuss applications and/or implications that might arise from this piece of research. [5 marks]**

If brief CBT works as well as traditional CBT, this has practical applications. It would save time for those taking part and for therapists delivering CBT • The wider spread of results in the brief CBT condition implies that this form of therapy may be more effective for some people than others. Clinicians need to assess patients carefully to decide who might be suitable and complete homework • Psychological problems can improve with the passage of time without treatment (spontaneous remission – suddenly getting better) meaning that there may be some merit in asking people to wait for treatment if their symptoms are not severe • An important implication is cost effectiveness. If brief CBT works as well as the longer traditional format, this would have a number of advantages including reducing costs for the NHS.

The psychologist wished to investigate which aspects of therapy were most useful. She decided to interview a sample of the patients. The purpose of the interviews was to encourage patients to talk freely and in detail about their experiences of therapy. Two of the questions used by the psychologist were:

● 'Please tell me about the most important aspects of the therapy which helped you to improve'

and

● 'Please tell me about any aspects of the therapy which were less helpful to you in dealing with

your symptoms'.

6.   **Explain how you would record the data from these interviews and your reason for choosing this method. [3 marks]**

Possible methods include: • Audio recording • Video recording (filming) • Making written notes during the interview.

Eg: 'I would choose to audio record the data using my phone (1 mark). This would be less intrusive than filming the patient (1 mark) so they would be more likely to agree to take part or be honest

7.   **Explain how you would analyse the qualitative data from the interviews. [5 marks]**

Content analysis; this method would involve identifying important categories from a sub sample of interview responses (for example references to homework or warmth in the therapist). The researchers would then work through the written data, counting the number of occurrences of each of the categories to produce quantitative data

Thematic analysis: This method would involve reading and rereading (familiarisation) the written transcripts carefully. Coding would involve looking for words which cropped up repeatedly in transcripts. These could then be combined to reduce the number of codes into three or four themes. The data would stay in qualitative format and would not be reduced to numbers.

In an observational study, 100 cars were fitted with video cameras to record the driver's behaviour. Two psychologists used content analysis to analyse the data from the films. They found that 75% of accidents involved a lack of attention by the driver. The most common distractions were using a hands-free phone or talking to a passenger. Other distractions included looking at the scenery, smoking, eating, personal grooming and trying to reach something within the car.

8. **What is content analysis? (2 marks)**

Content analysis is a technique for analysing qualitative data of various kinds. Data can be placed into categories and counted (quantitative) or can be analysed in themes (qualitative).

9. **Explain how the psychologists might have carried out content analysis to analyse the film clips of driver behaviour. (4 marks)**

The psychologist could have begun by watching some of the film clips of driver behaviour. • This would enable the psychologist to identify potential categories which emerged from the data of the different types of distractions seen in the film. • Such categories/themes might include: passenger distractions, gadget distractions, etc. • The psychologists would then have watched the films again and counted the number of examples which fell into each category to provide quantitative data.

10. **Explain how the two psychologists might have assessed the reliability of their content analysis. (3 marks)**

inter-rater reliability or test-retest reliability. - 'The two psychologists could watch the films separately and devise a set of categories. They could compare these and use categories they both agreed on. They could carry out content analysis of the films separately and compare their answers looking for agreement

The psychologists then designed an experiment to test the effects of using a hands-free phone on drivers' attention. They recruited a sample of 30 experienced police drivers and asked them to take part in two computer-simulated driving tests. Both tests involved watching a three-minute film of a road. Participants were instructed to click the mouse as quickly as possible, when a potential hazard (such as a car pulling out ahead) was spotted. Each participant completed two computer-simulated driving tests:

● **Test A, whilst chatting with one of the psychologists on a hands-free phone**

● **Test B, in silence, with no distractions.**

The order in which they completed the computer tests was counterbalanced.

11. **Explain why the psychologists chose to use a repeated measures design in this experiment. (3 marks)**

A repeated measures design was chosen in this experiment: • to remove the effects of individual differences in reaction times which would occur if an independent groups design was used • to avoid the potential difficulties involved in matching participants • to reduce the number of participants required for the experiment

12. **Identify one possible extraneous variable in this experiment. Explain how this variable may have influenced the results of this experiment. (3 marks)**

• the nature and content of the conversation with the psychologist on the hands-free phone • interaction between the sex of the psychologist and sex of participant which could influence the type of conversation • the number of hazards in the computer-based test, hence difficulty of the tests • the presence of the hands-free headset could have produced distraction

Example: The chat with the psychologist was not controlled (1 mark) so the difficulty or number of questions could have varied (2 marks). This would influence the DV as more or less attention would be required (3 marks).

13. Explain how one factor in this experiment might affect its external validity. (3 marks)

External validity refers to how far the findings of the experiment can be generalised to real-life situations. The most likely answer here is that the hazard perception test was done using a computer test which does not resemble real-life driving situations. (No noise, stress, etc.)

14. Explain one or more ethical issues that the psychologists should have considered in this experiment. (4 marks)

Protection of participants from harm whilst studying the effects of a hands-free phone on driving. Two key issues here are the use of a computer-based test with no risk attached and of an experienced sample of police drivers. • Informed consent: Participants should be given full information about the nature of both tasks before deciding whether or not to participate. • Debriefing: A full debriefing should take place at the end of the experiment. This should provide feedback on performance and allow participants to ask questions if they wish to. • Freedom to withdraw: Participants should be made aware of their freedom to withdraw before and during the experiment. They should be made aware of their right to withdraw their data after the experiment. • Confidentiality: Individuals should not be identified, but should retain anonymity (use of numbers or initials instead of names)

15. Write a set of standardised instructions that would be suitable to read out to participants, before they carry out Test A, chatting on a hands-free phone. (5 marks)

The standardised instructions should include the following information: a. You will take part in a simulated driving test which will last for three minutes. b. Your task will be to identify potential hazards on the road ahead. c. When you see a hazard, you should press the mouse button as quickly as possible. d. Whilst you are doing the test, I will chat to you on a mobile phone and I would like you to reply using the hands-free mobile phone headset. e. Do you have any questions? For full marks, the instructions should adopt an appropriate formal tone

The computer simulator measured two aspects of driver behaviour:

● **the number of hazards detected by each driver**

● **the time taken to respond to each hazard, in seconds.**

The mean scores for each of these measures is shown in Table 1.

Mean scores	Test A: with hands-free phone	Test B: in silence
Number of hazards detected	26.0	23.0
Reaction time in seconds	0.45	0.27

16. **The psychologists then used an inferential statistical test to assess whether there was a difference in the two conditions. Identify an appropriate statistical test to analyse the difference in the number of hazards detected in the two conditions of this experiment. Explain why this test of difference would be appropriate. (3 marks)**

Wilcoxon test: • A repeated measures design was used (1 mark) as drivers take part in both the hands-free phone and non-phone (silent) conditions (1 mark). • A repeated measures design was used (1 mark) and the data can be treated as ordinal (1 mark).

17. **Replication is one feature of the scientific method. The psychologists decided to replicate this experiment using a larger sample of 250 inexperienced drivers. Explain why replication of this study would be useful. (3 marks)**

Replication is an important tool in the scientific method. It allows scientists to check findings and ensure that they are robust. In this study, replication is important, as the original sample is small (30 people) and specific (experienced police drivers). For this reason, replication on a larger sample will be used to check if findings apply outside this specific group.

Some studies have suggested that there may be a relationship between intelligence and happiness. To investigate this claim, a psychologist used a standardised test to measure intelligence in a sample of 30 children aged 11 years, who were chosen from a local secondary school. He also asked the children to complete a self-report questionnaire designed to measure happiness. The score from the intelligence test was correlated with the score from the happiness questionnaire. The psychologist used a Spearman's rho test to analyse the data. He found that the correlation between intelligence and happiness at age 11 was +0.42.

18. **Write an operationalised non-directional hypothesis for this study. (2 marks)**

'There is a relationship between happiness scores on a questionnaire and intelligence test scores'.

19. **Identify an alternative method which could have been used to collect data about happiness in this study. Explain why this method might be better than using a questionnaire. (4 marks)**

An interview would be a more appropriate method than a questionnaire as it enables questions to be clarified and responses to be probed, thus overcoming the main disadvantages of questionnaires. Students could also make a case for the analysis of diaries/written materials as a way of collecting data about happiness. These would generally overcome the problems of social desirability and demand characteristics inherent in questionnaires. Students could also make a case for the use of observation.

20. What is meant by internal validity? (1 mark)

Internal validity refers to how accurately a test or measuring instrument measures what it says it measures

21. Describe how the internal validity of the happiness questionnaire could be assessed. (3 marks)

Concurrent validity involves assessing how closely the scores on the happiness questionnaire match a different measurement of happiness obtained from the same participants, for example from family/teacher reports • Content validity involves asking experts in the field to check the content of the questionnaire to see how accurately it measures happiness • Face validity is less rigorous and involves looking at the questions to see if they are genuinely asking about happiness.

22. A Spearman's rho test was used to analyse the data. Give two reasons why this test was used. (2 marks)

Study is looking for a correlation (relationship) • Suitable for pairs of scores • The data type obtained is ordinal, at least ordinal or interval level • Linear relationship between scores

23. The psychologist used a non-directional hypothesis. Using Table 1, state whether or not the correlation between intelligence and happiness at age 11 (+0.42) was significant. Explain your answer. (3 marks)

N (number of participants)	Level of significance for a two-tailed test	
	0.10	0.05
	Level of significance for a one-tailed test	
	0.05	0.025
29	0.312	0.368
30	0.306	0.362
31	0.301	0.356

Calculated r_s must equal or exceed the table (critical) value for significance at the level shown.

the obtained value of + 0.42 exceeds the critical value for a twotailed test (.362) for N=30. The results are therefore statistically significant ($p \leq 0.05$)

24. Five years later, the same young people were asked to complete the intelligence test and the happiness questionnaire for a second time. This time the correlation was −0.29. With reference to both correlation scores, outline what these findings seem to show about the link between intelligence and happiness. (4 marks)

At age 11, there is a significant positive correlation between happiness and intelligence, demonstrating that more intelligent children tend to be happier. • At age 16, the correlation is not statistically significant. Students may also make the point that there may be a weak tendency for more intelligent teenagers to be less happy at 16 years of age, although this is not statistically significant.

25. **The report was subjected to peer review before it was published in a journal. What is meant by peer review? (2 marks)**

Peer review is the process of subjecting a piece of research to independent scrutiny by other psychologists working in a similar field who consider the research in terms of its validity, significance and originality

26. **Explain why peer review is an important aspect of the scientific process. (4 marks)**

Peer reviewed research may be accepted, sent back for revisions or rejected. Peer review is an important part of the scientific process because: • It is difficult for authors and researchers to spot every mistake in a piece of work. Showing the work to others increases the probability that weaknesses will be identified and addressed. • It helps to prevent the dissemination of irrelevant findings, unwarranted claims, unacceptable interpretations, personal views and deliberate fraud. • Peer reviewers also judge the quality and the significance of the research in a wider context. • This process ensures that published research can be taken seriously because it has been independently scrutinised by fellow researchers.

27. **A psychology student was asked to design an investigation to see whether taking exercise could increase feelings of happiness. She proposed to do an experiment. She decided to recruit a sample of volunteers who had just joined a gym, by putting up a poster in the gym. She planned to carry out a short interview with each volunteer and to give each one a happiness score. She intended to interview the volunteers again after they had attended the gym for six weeks and to reassess their happiness score to see if it had changed. The psychology student's teacher identified a number of limitations of the proposed experiment. Explain one or more limitations of the student's proposal and suggest how the investigation could be improved. (10 marks)**

There are a number limitations of the proposal included in the stem. Some of the most obvious are as follows: • The independent variable (exercise) is not operationalised. There is no attempt to specify the amount of exercise taken, frequency or intensity. These could vary substantially. • The DV (happiness) is measured through an interview. Interviews are prone to demand characteristics and social desirability effects. Both of these could affect the validity of the measurement. Students may suggest using a happiness questionnaire to measure the DV. • As the student proposes to carry out the interviews herself, there is a likelihood of investigator effects. An independent interviewer could be used to reduce this. • There is a lack of a control group for comparison purposes. The experiment could be modified to use an independent group design, with a control group who do not undertake an exercise programme. • The use of a volunteer sample means that the study is unlikely to be representative. • Ethical issues – although a volunteer sample has been recruited, there is no mention of informed consent, confidentiality, debriefing etc. • Competence, the student is unlikely to have received training to carry out interviews of this nature.

28. **Explain what is meant by replicability. Why is replicability an important feature of science? (5 marks)**

Replicability is the ability to check and verify scientific information. Candidates could explain replicability as: • the ability to repeat the method to assess if similar findings are achieved • the ability to achieve similar findings

Replicability is an important part of the scientific process. Scientific method involves defining a problem and formulating a hypothesis which is tested with empirical research. Research findings are an important part of this process. If we wish to draw conclusions from research studies, the procedures and findings should be repeatable. Unrepeatable results may imply flaws or lack of control within the method used and are of limited use in theory construction.

A maths teacher wondered whether there was a relationship between mathematical ability and musical ability. She decided to test this out on the GCSE students in the school. From 210 students, she randomly selected 10 and gave each of them two tests. She used part of a GCSE exam paper to test their mathematical ability. The higher the mark, the better the mathematical ability. She could not find a musical ability test so she devised her own. She asked each student to sing a song of their choice. She then rated their performance on a scale of 1–10, where 1 is completely tuneless and 10 is in perfect tune.

29. **Suggest a suitable non-directional hypothesis for this study. (3 marks)**

'There is a correlation (relationship) between pupils' scores on a test of mathematical ability and pupils' scores on a test of musical ability'.

30. **Why might the measure of musical ability used by the teacher lack validity? (3 marks)**

The main issue is that the teacher has made up her own test: • This involved subjective judgement on the part of the teacher who rates the students' musical ability. Her judgement may not reflect real differences in musical ability and is likely to differ from other people's judgement and/or any absolute criteria for tunefulness. • Lack of reliability in rating musical ability would compromise the validity of the measure. • As the students can choose the song they will sing, the rating of ability could reflect the teacher liking/dislike of the song rather than the student's ability. • The rating may be invalid as the students selected songs which varied in difficulty so the tunefulness reflected the difficulty of the song not the students' ability. • Operationalising musical ability as tuneful singing is a very narrow measure. Someone can have musical ability such as playing an instrument which would not be reflected by this measure.

31. **Explain how the teacher could have checked the reliability of the mathematical ability test. (3 marks)**

In the case of the maths test candidates could refer to split half or test retest as methods of checking reliability. They could also refer to checking the reliability of scoring by using two separate markers for the test and comparing the scores.

32. **Explain why the teacher chose to use a random sample in this study. (2 marks)**

The teacher chose to use a random sample because it would probably be more representative of the whole GCSE group than if she had used an opportunity or volunteer sample. Candidates could also say that she had ready access to her target population making it convenient for her to select a random sample.

Student	Mathematical ability test score	Musical ability rating
1	10	10
2	2	9
3	9	3
4	6	6
5	3	9
6	10	2
7	2	1
8	1	8
9	8	4
10	4	7

33. **sketch a graph to show the data in Table 1. Give the graph an appropriate title and label the axes. (3 marks)**
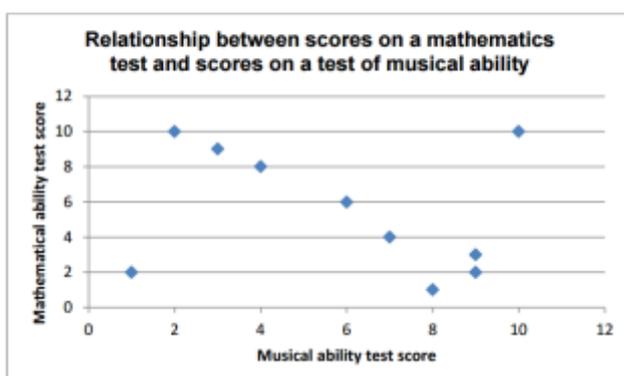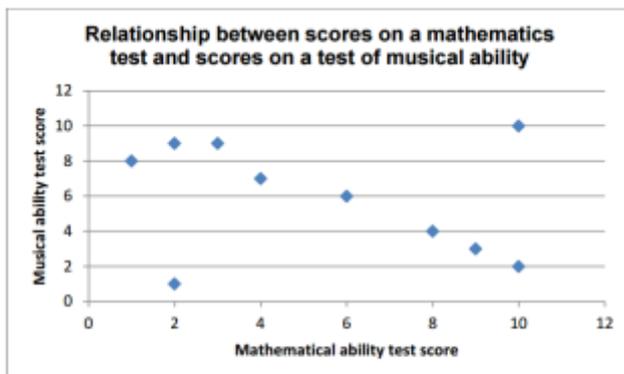
34. **Discuss what the data in Table 1 and the graph that you have sketched seem to show about the relationship between mathematical ability and musical ability. (3 marks)**

Likely points include: • The graph seems to show a negative correlation between mathematical and musical ability • This means that high scorers in mathematical ability tend to achieve low scores on

musical ability and vice versa • The presence of two strong outliers, means that the actual correlation is very weak and closer to zero. • comment on the small sample size which limits the conclusions that could be drawn • credit can be achieved for plausible interpretations of the strength of the correlation which are justified (ie looks moderate to strong or the outliers make it weak in practice) or those based on rough calculations (around -0.2).

35. **The teacher noticed that most of the students who were rated highly on musical ability were left-handed. The teacher is aware that her previous definition of musical ability lacked validity. Design a study to test whether there is a difference in the musical ability of left-handed students and right-handed students. You have access to a sixth form of 200 students. You should: identify the design that you would use, explain an appropriate sampling method and justify your choice, describe the procedure that you would use, including details of how you would assess musical ability, write a suitable debrief for these participants. (10 marks)**

Design – 1 mark

• Award 1 mark for identification of an appropriate design (independent measures or matched pairs).

Sampling – 2 marks

• Award 1 mark for explaining an appropriate sampling method and 1 further mark for justifying why this method would be appropriate. As left-handed people are less common in the population than right-handed people this needs to be addressed in the sampling method.

Procedure and assessment of musical ability – 4 marks

Award 1 mark for procedure, 1 mark for assessing musical ability and two further marks for elaboration of either or both of these.

• Description of the procedure eg each participant will be given a standardised musical ability test, participants should be tested within a controlled environment, with minimal noise or distraction.

• Students are required to suggest a plausible alternative method of assessing musical ability to the one in the stem (eg singing a short, novel phrase played on the piano). Further credit could be given for stating that the test should be identical for all students or for explaining how it will be assessed.

Debrief – 3 marks

• Award up to 3 marks for writing a debrief. This could include the aim of the study, thanking participants for taking part, asking if they have any questions, relevant ethical considerations.

• If this is not suitable to be read out to participants, maximum 1 mark

36. **In your answer book, draw a table to show how you would record your results. Identify an appropriate statistical test to analyse the data that you would collect. Justify your choice. (3 marks)**

Musical ability scores:

Participant number	Left handed	Right handed
1		
2		
3		

Award 1 mark for the identification of an appropriate statistical test for the proposed design.
Award 1 mark for one correct justification eg a test of difference, at least ordinal level data.