RES1 Research methods: variables, hypotheses, choosing a statistical test and the sign test

Total Marks

AQA 7182 · Calculator allowed · about 80 minutes

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

1 A researcher wants to investigate whether the amount of sleep a person has affects their performance on a short-term memory test. Participants are randomly allocated to either a Low Sleep group (less than 5 hours the previous night) or a Normal Sleep group (7 to 9 hours the previous night), then complete a memory test in which they try to recall as many words as possible from a 15-word list, after a short distractor task. The number of words correctly recalled is recorded for each participant.

(a) Identify the independent variable (IV) and the dependent variable (DV) in this study. **(2)**

(b) State the experimental design used in this study. **(1)**

(Total for Question 1 is 3 marks)

---

2 A hypothesis must be fully operationalised, and a study also needs a null hypothesis.

(a) Write a fully operationalised, directional (one-tailed) hypothesis for the study in Question 1. **(3)**

(b) Write the null hypothesis for this study. **(2)**

(Total for Question 2 is 5 marks)

---

3 Extraneous variables must be controlled, and the choice of experimental design has both strengths and limitations.

(a) Identify one extraneous variable that could affect the sleep and memory study, and explain how it could be controlled. **(3)**

(b) Explain one strength and one limitation of using an independent groups design, rather than a repeated measures design, for this study. **(3)**

(Total for Question 3 is 6 marks)

4 The table below can be used to help choose an appropriate inferential statistical test, based on the experimental design and the level of measurement of the data.

Design	Nominal data	Ordinal/interval data (not normally distributed)	Interval/ratio data (normally distributed)
Independent groups	Chi-Squared test of association	Mann-Whitney U test	Unrelated (independent) t-test
Repeated measures / matched pairs	Sign test	Wilcoxon signed-rank test	Related (paired) t-test
Correlation between two co-variables	not applicable	Spearman's rank correlation coefficient	Pearson's product-moment correlation coefficient

Use the table above to answer the three scenarios below.

(a) A researcher tests whether more participants correctly recall a word (yes or no) after drinking caffeine than after a placebo drink, using two separate, independent groups of participants. State which statistical test should be used, and justify your choice, referring to the design and the level of measurement. **(3)**

(b) A researcher correlates 15 participants' total scores on an anxiety questionnaire (ordinal data) with their exam grade (also treated as ordinal data). State which test should be used, and justify your choice. **(3)**

(c) A researcher compares the same 10 participants' reaction times before and after sleep deprivation. The reaction-time data is interval-level but is not normally distributed. State which test should be used, and justify your choice. **(3)**

(Total for Question 4 is 9 marks)

5 A psychologist tested whether a brief mindfulness intervention reduces self-reported anxiety. Twelve participants completed an anxiety questionnaire (scored 0 to 20, where a higher score means greater anxiety) before and after a two-week mindfulness intervention (a repeated measures design).

Participant	Anxiety score before	Anxiety score after
1	15	10
2	12	13
3	18	11
4	9	9
5	14	8
6	16	9
7	11	15
8	13	7
9	17	12
10	10	6
11	15	11
12	12	5

(a) State a suitable directional hypothesis for this study. **(2)**

(b) For each participant, state whether their anxiety score showed an increase (+), a decrease (-), or no change (0) after the intervention. Identify which participant(s), if any, should be excluded from the sign test, and state the value of n used in the test. **(3)**

(Total for Question 5 is 5 marks)

6 Using your working from Question 5, and the table of critical values for the sign test below, answer the following.

n (after excluding any tied/zero differences)	Critical value of S ($p < 0.05$, one-tailed)
5	0
6	0
7	0
8	1
9	1
10	1
11	2
12	2
13	3
14	3
15	3

For the sign test, the result is significant only if the calculated value of S is less than or equal to the critical value shown for that value of n.

(a) State the value of S (the sign test statistic) for the data in Question 5. **(2)**

(b) State the critical value of S for this value of n at $p < 0.05$ (one-tailed), and state whether the result is significant. Explain your decision. **(3)**

(Total for Question 6 is 5 marks)

---

7 Using the data from Question 5 (n = 11 participants included in the sign test).

(a) Calculate the percentage of the 11 participants included in the sign test who showed a decrease in anxiety score. Give your answer to 1 decimal place. **(3)**

(b) Explain why the sign test result being 'significant' is not the same thing as saying anxiety scores decreased for every participant. **(3)**

(Total for Question 7 is 6 marks)

---

8 Before conducting a study, researchers set a significance level (often $p < 0.05$) to decide whether to reject the null hypothesis.

(a) Explain what is meant by a 'Type I error' and a 'Type II error' in the context of a psychological study. **(4)**

(b) Explain why psychologists conventionally use a significance level of $p < 0.05$, rather than a stricter level such as $p < 0.01$. **(2)**

(Total for Question 8 is 6 marks)

9 The anxiety questionnaire used in the study in Question 5 is an example of a self-report method.

(a) Explain one strength and one limitation of using a questionnaire, rather than a structured interview, to measure anxiety in this study. **(4)**

(b) State one way response bias (such as social desirability bias) could affect the anxiety questionnaire results, and suggest a way to reduce it. **(2)**

(Total for Question 9 is 6 marks)

---

10 The psychologist wants to recruit a sample of adults from the general population for the mindfulness study.

(a) Identify a sampling method the researcher could use to recruit an unbiased sample, and describe how it would be carried out. **(3)**

(b) Explain one limitation of this sampling method. **(3)**

(Total for Question 10 is 6 marks)

---

11 The mindfulness study involves collecting personal information about participants' anxiety levels.

(a) Identify two ethical issues the researcher must address in this study, and explain how each could be dealt with. **(4)**

(b) Explain why debriefing is particularly important in a study measuring anxiety. **(2)**

(Total for Question 11 is 6 marks)

Mark scheme · RES1 Research methods: variables, hypotheses, choosing a statistical test and the sign test

Question 1

- (a) **B1** IV: amount of sleep the previous night (Low Sleep group versus Normal Sleep group) oe
- (a) **B1** DV: number of words correctly recalled from the 15-word list oe
- **(a) Answer: IV: amount of sleep (Low Sleep vs Normal Sleep group). DV: number of words correctly recalled.**
- (b) **B1** independent groups (independent measures) design oe
- **(b) Answer: Independent groups (independent measures) design.**

Question 2

- (a) **M1** states a direction, e.g. participants in the Low Sleep group will recall fewer words than participants in the Normal Sleep group oe
- (a) **A1** operationalises the IV: specifies the sleep amounts defining each group (less than 5 hours vs 7 to 9 hours) oe
- (a) **A1** operationalises the DV: specifies the measure, e.g. number of words correctly recalled from the 15-word list oe
- **(a) Answer: Participants who have less than 5 hours' sleep the previous night will correctly recall significantly fewer words from a 15-word list than participants who have 7 to 9 hours' sleep the previous night.**
- (b) **B1** states there will be no significant difference between the two groups' recall oe
- (b) **B1** attributes any difference found to chance, rather than the manipulated IV oe
- **(b) Answer: There will be no significant difference in the number of words correctly recalled between the Low Sleep group and the Normal Sleep group; any difference found will be due to chance.**

Question 3

- (a) **M1** identifies a plausible extraneous variable, e.g. time of day the memory test is taken, or participants' caffeine intake oe
- (a) **A1** suggests a suitable control, e.g. testing all participants at the same time of day oe
- (a) **A1** explains why this controls the variable: it cannot then differ systematically between the two groups and confound the results oe
- **(a) Answer: E.g. time of day the test is taken; controlled by testing all participants at the same time of day, so it cannot confound the results.**
- (b) **M1** strength: no order effects (e.g. practice or fatigue from doing the memory test twice), since each participant only takes part in one condition oe
- (b) **A1** limitation: participant variables (e.g. natural differences in memory ability) are not controlled, since different people make up each group oe
- (b) **A1** explains that this could confound the results if, by chance, one group happens to contain participants with generally better memories oe
- **(b) Answer: Strength: no order effects, since each participant does the test only once. Limitation: participant variables are not controlled between the two different groups of people.**

Question 4

- (a) **B1** Chi-Squared test of association cao
- (a) **M1** justifies using an independent groups design oe
- (a) **A1** justifies using nominal/categorical data, correct or incorrect counts oe
- **(a) Answer: Chi-Squared test of association - an independent groups design with nominal (categorical) data.**
- (b) **B1** Spearman's rank correlation coefficient cao
- (b) **M1** justifies using a correlation between two co-variables oe
- (b) **A1** justifies using ordinal-level data oe
- **(b) Answer: Spearman's rank correlation coefficient - a correlation using ordinal data.**
- (c) **B1** Wilcoxon signed-rank test cao
- (c) **M1** justifies using a repeated measures design, the same participants tested twice oe
- (c) **A1** justifies using interval data that is not normally distributed oe
- **(c) Answer: Wilcoxon signed-rank test - a repeated measures design with interval data that is not normally distributed.**

Question 5

- (a) **M1** states a direction: anxiety scores will be lower after the intervention than before oe
- (a) **A1** operationalises both measures, referring to the 0 to 20 anxiety questionnaire score taken before and after the intervention oe
- **(a) Answer: Participants' scores on the 0 to 20 anxiety questionnaire will be significantly lower after the two-week mindfulness intervention than before it.**
- (b) **B1** correctly identifies the signs: participant 1 -, 2 +, 3 -, 4 0, 5 -, 6 -, 7 +, 8 -, 9 -, 10 -, 11 -, 12 - oe
- (b) **B1** identifies participant 4 as excluded, since their score showed no change (a zero difference) oe
- (b) **B1** states n = 11 cao
- **(b) Answer: Signs: 1 -, 2 +, 3 -, 4 (excluded, no change), 5 -, 6 -, 7 +, 8 -, 9 -, 10 -, 11 -, 12 -. Participant 4 is excluded; n = 11.**

Question 6

- (a) **M1** identifies the less frequent sign: the '+' sign occurs for participants 2 and 7 only, while '-' occurs for the other 9 included participants oe
- (a) **A1** S = 2 cao
- **(a) Answer: S = 2.**
- (b) **B1** states the decision rule: for the sign test, the result is significant only if the calculated value is less than or equal to the critical value oe
- (b) **M1** identifies the critical value for n = 11 from the table as 2 oe
- (b) **A1** since the calculated value of S (2) is less than or equal to the critical value (2), the result is significant at $p < 0.05$ (one-tailed); the null hypothesis is rejected oe
- **(b) Answer: Critical value = 2. Since S (2) is less than or equal to 2, the result is significant at $p < 0.05$ (one-tailed); the null hypothesis is rejected.**

Question 7

- (a) **M1** identifies that 9 of the 11 included participants showed a decrease oe
- (a) **M1** sets up the calculation: (9 / 11) x 100 oe
- (a) **A1** = 81.8% (1dp) awrt 81.8
- **(a) Answer: 81.8% (1 decimal place).**
- (b) **M1** statistical significance means the pattern found is unlikely to have occurred by chance, not that the effect is large or seen in every case oe
- (b) **A1** in this data, 2 of the 11 included participants actually showed an increase, not a decrease, in anxiety oe
- (b) **A1** so a significant result can still occur even though the effect was not seen in every single participant oe
- **(b) Answer: Significance means the pattern is unlikely to be due to chance, not that every participant showed the effect; here, 2 of the 11 participants actually increased.**

Question 8

- (a) **M1** a Type I error is rejecting the null hypothesis (concluding there is a significant effect) when the null hypothesis is actually true, a 'false positive' oe
- (a) **A1** e.g. concluding the mindfulness intervention worked when it actually had no real effect oe
- (a) **M1** a Type II error is failing to reject the null hypothesis (concluding there is no significant effect) when the null hypothesis is actually false, a 'false negative' oe
- (a) **A1** e.g. concluding the mindfulness intervention did not work when it actually did have a real effect oe
- **(a) Answer: Type I error: wrongly concluding there is a significant effect when there is not. Type II error: wrongly concluding there is no significant effect when there is one.**
- (b) **M1** $p < 0.05$ balances the risk of a Type I error against the risk of a Type II error oe
- (b) **A1** a stricter level such as $p < 0.01$ reduces the risk of a Type I error but increases the risk of missing a genuine effect (a Type II error), so $p < 0.05$ is a reasonable, conventional compromise oe
- **(b) Answer: $p < 0.05$ is a conventional balance between the risk of a Type I error and the risk of a Type II error.**

Question 9

- (a) **M1** strength: a questionnaire can be given to many participants at once, and is quick/cheap to administer and score oe
- (a) **A1** explains a benefit of this, e.g. allowing a larger sample to be tested, increasing the reliability/generalisability of the findings oe
- (a) **M1** limitation: participants may not always answer honestly, e.g. due to social desirability bias, wanting to appear less anxious than they really are oe
- (a) **A1** explains a consequence, e.g. this could reduce the validity of the anxiety scores collected oe
- **(a) Answer: Strength: quick and cheap to administer to many participants at once. Limitation: participants may not answer honestly (e.g. social desirability bias), reducing validity.**
- (b) **B1** participants may under-report their true level of anxiety, to appear more calm/socially acceptable oe
- (b) **B1** suggests a way to reduce this, e.g. assuring participants that their answers are anonymous/confidential oe
- **(b) Answer: Participants may under-report anxiety to appear more socially acceptable; assuring anonymity/confidentiality can help reduce this.**

Question 10

- (a) **M1** identifies a suitable probability sampling method, e.g. random sampling oe
- (a) **A1** describes the method, e.g. obtaining a complete list of the target population and using a random method, such as a random number generator, to select participants from it oe
- (a) **A1** explains that every member of the target population has an equal chance of being selected oe
- **(a) Answer: Random sampling: obtain a full list of the target population and use a random method (e.g. a random number generator) to select participants, giving everyone an equal chance of selection.**
- (b) **M1** identifies a genuine limitation of random sampling oe
- (b) **A1** e.g. obtaining a complete, accurate list of the whole target population can be difficult or impossible in practice oe
- (b) **A1** and/or selected individuals may refuse to take part, meaning the final sample may still not be fully representative despite being randomly selected oe
- **(b) Answer: A complete list of the population may be hard to obtain, and selected people may refuse to take part, so the final sample may still not be fully representative.**

Question 11

- (a) **M1** identifies a genuine ethical issue, e.g. confidentiality oe
- (a) **A1** explains how it is addressed, e.g. storing questionnaire data anonymously, using participant numbers rather than names oe
- (a) **M1** identifies a second genuine ethical issue, e.g. informed consent oe
- (a) **A1** explains how it is addressed, e.g. giving participants full information about the study before they agree to take part, and confirming their right to withdraw at any time oe
- **(a) Answer: E.g. confidentiality, addressed by anonymising data; and informed consent, addressed by fully briefing participants and confirming their right to withdraw.**
- (b) **B1** debriefing allows the researcher to check on participants' wellbeing/psychological state after completing the study oe
- (b) **B1** this is especially important here because the topic, anxiety, is personally sensitive, and participants may need reassurance, information, or signposting to support if the study has raised any concerns oe
- **(b) Answer: Debriefing lets the researcher check participants' wellbeing and offer support/information, which matters more given the personally sensitive topic of anxiety.**