7.13 Reliability and Validity in Psychological Research: Test Pack

AQA 7182 · about 20 minutes

Total Marks

Name: _____ Date: ____ / ____ / _____

Answer ALL questions.

1 Define the term 'reliability' as it applies to psychological measurement in research, and give one brief example of reliable measurement in a psychology study context.

(Total for Question 1 is 2 marks)

2 Suggest one practical way to improve face validity of a short depression screening questionnaire before piloting it with participants.

(Total for Question 2 is 2 marks)

3 A researcher uses two observers to record instances of helping behaviour in a school playground. Observer A records 18 helping acts in an hour and Observer B records 12. Outline one reliability issue this raises and suggest one specific improvement to increase inter-observer reliability.

(Total for Question 3 is 4 marks)

© Revision Library | revisionlibrary.co.uk

4 Researchers develop a new short questionnaire intended to measure social anxiety and want to demonstrate concurrent validity. They administer the new questionnaire and an established clinical social anxiety inventory at the same session to 60 participants. Explain how concurrent validity is assessed in this situation and give one improvement the researchers could make to strengthen their evidence for validity.

(Total for Question 4 is 3 marks)

---

5 Outline two ways standardisation can improve both reliability and internal validity in a lab experiment that measures reaction time to a visual stimulus.

(Total for Question 5 is 4 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 7.13 Reliability and Validity in Psychological Research: Test Pack

Question 1

definition: consistency of a measure, producing similar results on repeated occasions or between observers oe **B1**

example: a standardised memory test that gives similar scores to the same participant when repeated under the same conditions oe **B1**

Answer: Reliability is the consistency of a measure over time or between observers. Example: a standardised vocabulary test that yields similar scores for the same person when taken twice under the same conditions.

Question 2

suggests a practical step, e.g. ask experts/clinicians to review items or run a small panel of laypeople to check whether items appear to measure depression oe **B1**

explains how this helps: expert or lay review ensures items look appropriate and understandable, increasing apparent relevance and clarity oe **B1**

Answer: Ask mental health clinicians and a small panel of lay users to review the items for relevance and clarity. Their feedback ensures the items appear to measure depression and are understandable, improving face validity.

Question 3

identifies a reliability issue: low inter-observer agreement, observers producing different counts oe **M1**

explains why this is a problem: it threatens the consistency of measurement and may reflect unclear operational definitions oe **A1**

suggests a specific improvement, e.g. provide clear operational definitions and train observers with practice coding and trial sessions oe **M1**

explains how this would help: training and agreed definitions increase agreement and make counts more consistent across observers oe **A1**

Answer: Issue: low inter-observer agreement, shown by observers recording very different counts, implying inconsistent measurement. Improvement: give clear operational definitions of 'helping' and run observer training with practice coding until agreement is high; this standardises judgements and increases inter-observer reliability.

Question 4

explains assessment: concurrent validity is assessed by correlating scores from the new measure with scores from a well-established measure administered at the same time oe **M1**

states the expected pattern: a strong positive correlation would indicate good concurrent validity oe **B1**

gives an improvement, e.g. increase sample diversity/size or use multiple established measures and a range of clinical and non-clinical participants to check convergent validity and generalisability oe **A1**

Answer: Concurrent validity is assessed by correlating scores on the new questionnaire with scores on the established social anxiety inventory administered at the same time; a strong positive correlation supports concurrent validity. To strengthen evidence, researchers could increase sample size and include a wider range of participants, and compare the new measure with several established scales to check convergent validity and generalisability.

Question 5

identifies first way: provide standardised instructions and practice trials for all participants oe **M1**

explains effect: this reduces variation due to misunderstanding and increases consistency of measurement, improving reliability and reducing procedural confounds that harm internal validity oe **A1**

identifies second way: control environmental factors such as lighting, screen distance and stimulus timing using the same equipment and software for all participants oe **M1**

explains effect: controlling these factors reduces extraneous variation in reaction times and helps ensure differences are due to the manipulated variable, improving internal validity and reliability oe **A1**

Answer: Standardise instructions and include practice trials so every participant understands the task, reducing random variation and increasing reliability. Also control environmental and equipment factors, such as lighting, viewing distance and stimulus timing, so differences in reaction time are not due to extraneous causes, improving internal validity and consistency.

© Revision Library | revisionlibrary.co.uk