

6.1 Measurement, methods, statistics and history: evidence and transfer

Name: _____ Class: _____ Date: _____

Total: 8 marks

KEY WORDS reliability • validity

Objective

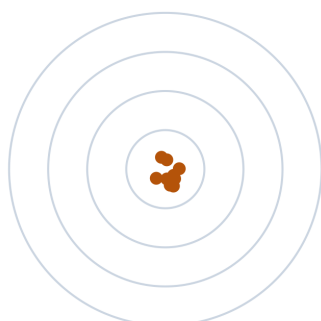
Distinguish reliability, validity and operational definitions; Analyse experiments, sampling, ethics and statistical inference; Recognise foundational schools and their explanatory commitments. Original practice and authored marking guidance.

1 Key knowledge and worked examples

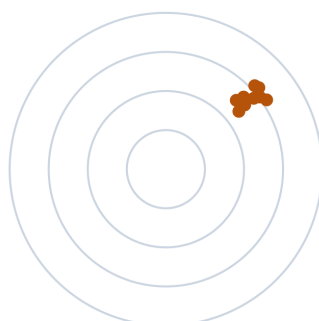
reliability : Consistency of measurement. **validity** : Evidence for an intended measurement interpretation.

Define a construct, state its observable measure, then ask whether the measure supports the intended interpretation. Random sampling concerns who enters a study; random assignment concerns which condition participants receive. A causal conclusion needs a credible comparison as well as measurement and ethical safeguards. Historical schools differ in the questions and methods they prioritize. See MM.1, MM.2 and HT.1 before using statistical or historical labels as explanations.

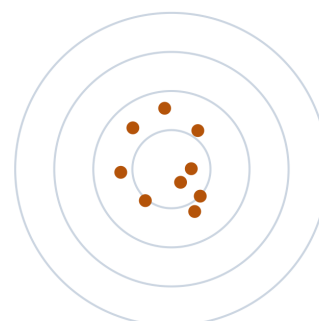
Reliable + valid



Reliable, not valid



Unreliable



Worked case. Volunteers are randomly assigned to two memory tasks. Does this make the sample representative of all adults?

No. Assignment improves comparability of conditions within the recruited sample; it

does not undo volunteer selection. Generalization requires attention to recruitment, eligibility and population differences. Describe random sampling and assignment separately.

2 Practice

2.1 A stable wrong measurement shows: [1]

A. reliability without demonstrated validity B. perfect validity C. no consistency

2.2 Random assignment primarily supports: [1]

A. causal group comparison B. representative population sampling C. automatic external validity

2.3 Revisit the worked case. State one conclusion its evidence supports and one stronger conclusion it does not support. Explain the difference. [2]

3 Exam-style questions

3.1 A report finds $p=.03$ and concludes that its theory has a 97% probability of being true and a large effect. Correct both inferences. Give the interpretation, supporting reasoning, a relevant alternative or limitation, and an appropriate check. [4]

Solutions

2.1 Answer: A. Repeatability alone does not establish accuracy or construct fit. The other options do not satisfy the stated defining conditions.

2.2 Answer: A. It balances group differences in expectation. Compare the exact process, design or evidence, rather than an associated keyword.

2.3 Use the worked case evidence to support a bounded claim and explain why a causal, universal or stronger claim needs additional evidence. The claim must fit the specific design and observed process.

3.1 Under the specified null model and assumptions, the probability of results at least as extreme as the observed result is .03. It is not a posterior probability for the theory. Effect size and uncertainty must be reported separately; even a small effect can have a small p-value in a large sample. Design bias is not removed by significance. Award B1 each (authored) for the interpretation, its reasoning, a relevant limitation/alternative and an appropriate check, up to four points. Accept a different justified answer satisfying these demands.