

# Measurement, methods, statistics and history

Supported GRE teaching and original transfer practice

## Identify the evidence

### Identify the evidence

Distinguish reliability, validity and operational definitions

Read MM.1, MM.2 and HT.1 for measurement, design and historical claims.

# Concepts 1

## Concepts 1

Reliability concerns consistency; validity concerns evidence supporting a particular interpretation or use.

Operational definitions specify how a construct is measured. A questionnaire label does not establish construct validity.

## Concepts 2

### Concepts 2

Random assignment improves comparisons; random sampling supports population generalisation. Ethics constrains both design and interpretation.

A p-value is a probability of data at least as incompatible under the null model, not the probability that the null is true. History questions distinguish explanatory schools and their methods.

## Reason through the worked case

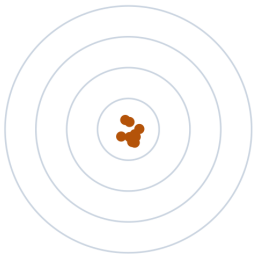
### Reason through the worked case

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

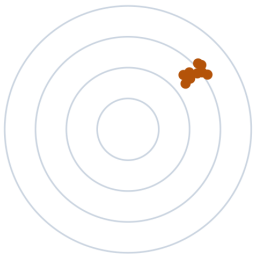
No. Assignment improves comparability of conditions within the recruited sample; it does not undo volunteer selection. Generalization requires attention to recruitment, eligibility and population differences. Describe random sampling and assignment separately.

## Read the reviewed representation

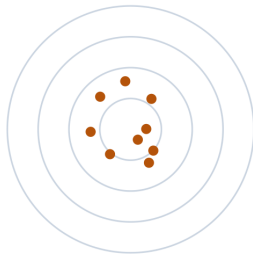
Reliable + valid



Reliable, not valid



Unreliable



Original schematic for Psychometrics and test construction. Illustrative relationships and numbers are not research results.

## Attempt an unfamiliar case

### Attempt an unfamiliar case

A report finds  $p=.03$  and concludes that its theory has a 97% probability of being true and a large effect. Correct both inferences.

Pause and write the explanation before continuing.

## Check the transfer argument

### Check the transfer argument

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.

Accept a different explanation when its evidence and limits are justified.

## Limit the conclusion

### Limit the conclusion

Statistical significance alone does not tell you effect size, practical value or absence of design bias.

Original classroom cases are not official questions or personal clinical advice.

## Check yourself

### Independent retrieval

Random sampling is used but allocation is not random. Which inference is strengthened and which causal problem remains?

Write your reasoning before comparing it with the methods and limits taught in this lesson.