Exercise walk-through

11.1

Research designs and statistical interpretation: evidence and transfer

GRE · GRE Subject Test · GRE Psychology

Designs

- A true experiment manipulates an independent variable and assigns participants to conditions at random.
- Between-subjects: different people in each condition. Within-subjects: the same people in every condition, which needs **counterbalancing** of the order.
- A quasi-experiment compares groups that were not randomly assigned, such as two existing classes.
- A correlational study measures without manipulating. Correlation does not show causation or which came first.

Reading statistics

- A p-value is the probability of results at least this extreme if the null hypothesis and the model's assumptions were true. It is not the probability that the null hypothesis is true.
- Statistical significance is not effect size. A tiny effect can be significant in a huge sample.
- Type I error: rejecting a true null hypothesis. Type II error: failing to reject a false one.
- A mean of 26 with most students aged 25 or younger suggests right skew, but does not prove the shape.

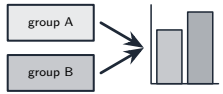
Key terms

- **random assignment**: Allocating participants to conditions by chance.
- **counterbalancing**: Varying the order of conditions across participants so that order effects cancel.
- **internal validity**: Confidence that the manipulated variable caused the observed difference.
- **external validity**: Confidence that results generalise beyond the study.

Exercises 2.1–2.2

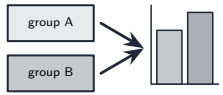
2.1 Each participant learns a list using exactly one of three rehearsal strategies. The design is: [1]

A. within-subjects B. between-subjects C. correlational D. longitudinal E. quasi-experiment



2.2 Counts of students choosing each of three majors, broken down by gender, are best analysed with: [1]

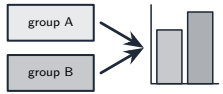
A. an independent-samples t-test B. one-way analysis of variance C. a chi-square test D. a Pearson correlation E. linear regression



Exercises 2.1·2.2

2.1 Each participant learns a list using exactly one of three rehearsal strategies. The design is: [1]

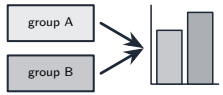
A. within-subjects B. between-subjects C. correlational D. longitudinal E. quasi-experiment



2.1 Answer: B. Different participants serve in each condition, with the strategy manipulated between them. The other options do not satisfy the stated defining conditions.

2.2 Counts of students choosing each of three majors, broken down by gender, are best analysed with: [1]

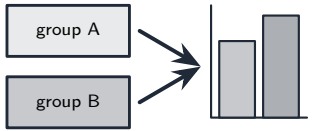
A. an independent-samples t-test B. one-way analysis of variance C. a chi-square test D. a Pearson correlation E. linear regression



2.2 Answer: C. The data are categorical counts in a 2×3 table, $\mathrm{df}=(2-1)(3-1)=2$. Compare the exact process, design or evidence, rather than an associated keyword.

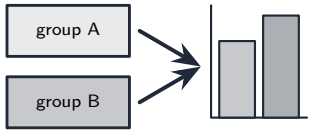
Exercise 2.3

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



Exercise 2.3

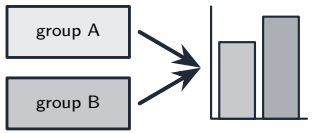
2.3 Revisit the worked case. State one conclusion its evidence supports and one stronger conclusion it does not support. Explain the difference. [2]



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.

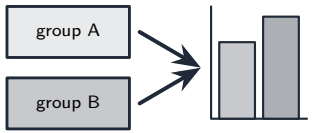
Exercise 3.1

3.1 The same 24 people complete two tasks. Another analyst treats the 48 scores as independent. Explain the problem and a suitable analysis family. Give the interpretation, supporting reasoning, a relevant alternative or limitation, and an appropriate check. [4]



Exercise 3.1

3.1 The same 24 people complete two tasks. Another analyst treats the 48 scores as independent. Explain the problem and a suitable analysis family. Give the interpretation, supporting reasoning, a relevant alternative or limitation, and an appropriate check. [4]



3.1 Scores within a person are related. Ignoring pairing uses an inappropriate independence assumption and loses information about within-person differences. For suitable quantitative difference data and assumptions, a paired t-test may fit; a justified nonparametric alternative may be appropriate otherwise. Randomized or counterbalanced order also addresses sequence effects. 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.