Research designs and statistical interpretation

Supported GRE teaching and original transfer practice

Research designs and statistical interpretation

- Tell experiments from quasi-experiments and correlational studies.
- Separate **internal validity** from **external validity**.
- Choose an analysis that fits the design.

Research designs and statistical interpretation (2)

- Read main effects and the **interaction effect 交互作用**.
- Interpret p-values, effect sizes and errors correctly.

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 平衡抵消** 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.

Two kinds of randomness

- **Random assignment** supports **internal validity** 内部效度: confidence that the manipulation caused the difference.
- Random sampling supports **external validity** 外部效度: confidence that results apply to the wider population.
- A study can have one without the other.

Matching the analysis

- Two independent groups, quantitative scores: an independent-samples t-test.
- The same people measured twice: a paired t-test.
- Several factors: analysis of variance (ANOVA).

Matching the analysis (2)

- Counts in categories: chi-square, if expected counts are large enough.
- Associations: correlation or regression, which still do not show causation.

Main effects and interactions

- A main effect is the effect of one factor averaged over the other.
- An interaction means the effect of one factor depends on the level of the other. On a graph of means, the lines are not parallel.
- An interaction can exist even when every main effect is zero.

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.

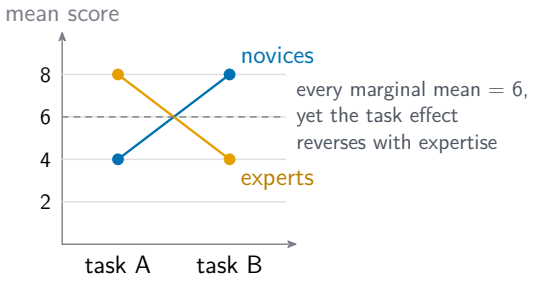
exercises.lol exercises.com.cn

Worked example: a crossover

- Novices average $(4 + 8)/2 = 6$; experts average 6. Each task also averages 6.
- So there is no main effect of expertise or of task.
- Yet the task effect reverses with expertise: a complete interaction.

exercises.lol exercises.com.cn

Figure: Worked example: a crossover

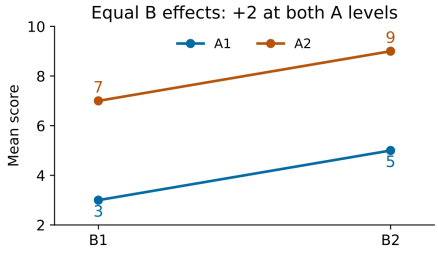


- Equal marginal means can hide a full interaction.
- Match each part of the diagram to the text.

Worked example: parallel lines

- B adds 2 points at both levels of A, and A adds 4 points at both levels of B.
- The lines are parallel: two main effects and no interaction.
- If the same 24 people gave all the scores, the 48 scores are not independent. A repeated-measures analysis is needed.

Figure: Worked example: parallel lines



Invented balanced examples; no significance test is shown.

- Parallel lines mean B adds +2 at both levels of A: main effects and no interaction
- Read the axes and units first.

Common errors

- Drawing causal conclusions from a correlational study.
- Confusing **random assignment** with random sampling.
- Treating paired scores as independent.

Common errors (2)

- Reporting an interaction as a main effect of one factor.
- Reading p as the probability that the null hypothesis is true.

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.

Check yourself

- A study compares two existing schools. Is it a true experiment?
- Which kind of randomness supports causal conclusions?
- On a graph of means the lines cross. What does this suggest?