

Research designs and statistical interpretation

GRE · GRE Subject Test · GRE Psychology

Scope and prerequisites

Supported GRE topic 11 (MM.2). Use the overview measurement lesson VI when interpreting research evidence. Existing official questions are traced in `bank_review_form.yaml`; original tasks below are not ETS items.

Concepts and method

interaction effect : The dependence of one factor's effect on the level of another factor. **random assignment** : Allocating participants to conditions by chance to support causal inference.

A true experiment manipulates an independent variable with random assignment of participants to conditions; between-subjects designs give different people to each condition, within-subjects designs measure the same people in every condition and need counterbalancing. A quasi-experiment lacks full assignment, and a correlational study does not manipulate an independent variable. Correlation alone does not establish causation or temporal order. Random assignment supports internal validity (causal inference); random sampling supports external validity (generalisability). They answer different questions.

Match analysis to design, measurement scale and assumptions. Two independent quantitative groups may suit an independent-samples t-test; paired data require a paired analysis. Factorial designs often use analysis of variance; categorical counts may suit a chi-square test when expected-count assumptions hold. Correlation or regression examines specified associations; neither alone establishes causation.

A main effect averages one factor across the other factor. An interaction means one effect depends on the other factor. Nonparallel sample means suggest an interaction pattern, but variability and the design are needed for population inference. Interactions can exist when averaged main effects cancel; lack of significance is not proof of no effect.

Mean age 26 with most students 25 or younger is compatible with right skew, but does not uniquely determine distribution shape. Statistical significance is not effect size. A p-value is a probability of results at least as extreme under the specified null model and assumptions, not the probability the null is true. Type I and Type II errors depend on design and decision thresholds.

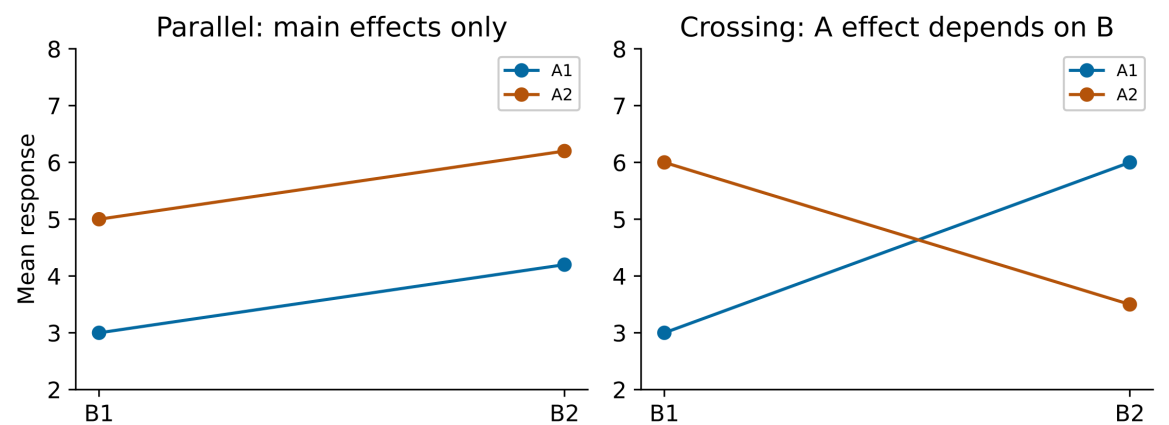
Choose a test from the design, scale and assumptions, not merely the number of groups. Paired observations require a paired or repeated-measures analysis; independent groups require an independent-group analysis. An interaction can be meaningful

even if averaged main effects cancel. Statistical significance is evidence under a model, not a condition for describing an observed difference. Report estimates and uncertainty and qualify population claims.

Worked reasoning

Fictional mean scores are: task A, novices 4 and experts 8; task B, novices 8 and experts 4. Compute marginal means and describe the pattern.

Novice marginal mean= $(4+8)/2=6$; expert marginal mean= $(8+4)/2=6$. Task A marginal mean= $(4+8)/2=6$; task B marginal mean= $(8+4)/2=6$. The task contrast reverses by expertise, an interaction pattern despite equal marginal means. Means alone do not establish a population interaction; variability, design and inference are needed.



Independent transfer

The same 24 people complete two tasks. Another analyst treats the 48 scores as independent. Explain the problem and a suitable analysis family.

Attempt independently, then use the matching skill sheet solution.

Review and source limits

Avoid this error: Drawing a causal conclusion from a correlational design, or reporting a significant interaction as if it were a main effect of one factor. Teaching scope comes from teaching.yaml and the source/key/crop review in bank_review_form.yaml. Detailed current diagnostic criteria require an authenticated manual; neither historical ETS practice nor this educational reference certifies them.