Research & Analysis / Academic Research

Evaluate study design, sample size, statistical power, and potential bias in academic papers.
Difficulty: Advanced
Model: GPT-4 / Claude / Gemini
Use Case: Peer Review, Paper Critique, Methods Sections
Updated: May 2026
Why This Prompt Exists
Not all published research is good research — but spotting methodological flaws takes expertise.

You get:

- citing papers with fatal methodological flaws
- building your research on weak foundations
- missing obvious confounds in studies you review
- peer reviews that focus on writing instead of methods
- wasted time on studies that can’t support their claims

But methodological critique follows checklists:

- design type: experimental, quasi-experimental, correlational, qualitative
- sample: size, power, representativeness, attrition
- measurement: validity, reliability, common method bias
- analysis: assumptions, multiple comparisons, p-hacking signs
- causality: randomization, temporal precedence, alternative explanations

Without systematic analysis, you trust the peer review process too much.

This prompt evaluates a paper’s methodology against standard quality criteria.

The Prompt

```
Assume the role of a methodological statistician who evaluates
research quality.

Your task is to analyze a paper's methodology for strengths and flaws.

Generate:

1. STUDY DESIGN ASSESSMENT
   - Design type
   - Internal validity (high/medium/low)
   - External validity (high/medium/low)
   - Key threats (selection, history, maturation, testing,
instrumentation)

2. SAMPLE ASSESSMENT
   - Sample size (underpowered / adequate / overpowered)
   - Power analysis reported? (Y/N — and if Y, what effect size
assumed)
   - Sampling method (random, convenience, purposive, snowball)
   - Attrition rate and handling

3. MEASUREMENT ASSESSMENT
   - Construct validity (do measures capture the construct?)
   - Reliability reported? (Cronbach's alpha, inter-rater)
   - Common method bias risk

4. ANALYSIS ASSESSMENT
```

- Statistical assumptions checked? (normality, homogeneity, independence)
- Multiple comparison corrections? (Bonferroni, FDR, none)
- Signs of p-hacking (just barely significant, selective reporting)

5. OVERALL JUDGMENT
- Can the paper's conclusions be trusted? (Yes / With caveats / No)
- Most serious limitation (one sentence)
- One question for the authors

INPUTS:

Paper (methods section + results):
[PASTE OR UPLOAD]

Field/discipline:
[E.G., "Psychology", "Epidemiology", "Economics"]

Your expertise level:
[BEGINNER / GRADUATE / EXPERT]

RULES:
- Be specific — "sample size too small" is less useful than "n=30 detects only r>.5, but they claim r=.2"
- Flag what's missing, not just what's wrong
- Distinguish between fatal flaws (reject paper) and minor issues (fix with caveats)
- Note when a flaw is common in the field (not an excuse, but context)

How To Use It

- Run this before citing a paper in your own work — especially if the finding is surprising.
- Use as a peer review checklist before submitting your own papers.
- For course readings, run this to prepare class discussion questions.
- Pay closest attention to the “overall judgment” — some papers aren’t worth your time.
- Save the “questions for authors” — that’s your starting point for a commentary or replication.

Example Input

Paper (methods section + results):
“We surveyed 120 undergraduate psychology students (72 female, mean age 19.4). Participants completed our 5-item measure of grit (α = .82) and a single-item measure of academic success (self-reported GPA). Correlations were computed. No power analysis reported. p-values below .05 considered significant.”

Field/discipline:
Psychology

Your expertise level:
Graduate

Why It Works
Most paper reading stops at the abstract and discussion — methodology gets skimmed.

This framework improves outcomes by forcing:

- design assessment (internal vs. external validity trade-offs)
- sample evaluation (underpowered studies waste everyone’s time)
- measurement critique (bad measures = bad conclusions)

- p-hacking detection (statistical fraud is real)
- clear judgment (cite with confidence or walk away)

Great methodology analysis doesn't tear papers down — it tells you which ones to trust.

Build Better AI Systems

Subscribe for advanced prompt engineering, AI coding tools, debugging frameworks, and practical strategies for developers and engineers.

Carefully engineered prompts for people doing real work.

Share this:

See also Abstract to Full Paper Expander