

5.2 Choose a method that measures the question

Name: _____ Class: _____ Date: _____

Total: **24 marks****KEY WORDS** survey 问卷调查 • sample 样本 • random assignment 随机分组 • pilot 试测

Objective

Choose a method that measures the question. This is original GAC033-aligned practice, not an official assessment. All people, cards, records and instructions are fictional; no main study has been conducted.

1 Key knowledge and worked examples

■ Key knowledge: preference and accuracy are different outcomes

A **survey** 问卷调查 can ask about experience or preference; it does not by itself show accurate sorting. Match the method to the outcome. The **sample** 样本 is the participating group. **Random assignment** 随机分组 allocates conditions within that group; it is different from randomly sampling a whole school.

A **pilot** 试测 checks whether a procedure is understandable and workable. If instructions change after the pilot, keep its sheets separate from the main comparison. A plan or a pilot is not a completed main study.

■ Source A: proposal and task record

The fictional question compares completely correct four-item paper sorting sheets under layouts A and B. The proposed sample is twenty-four willing peers from one class: twelve assigned to each layout by shuffled slips. Each sees one layout and has the same words, item choices, two-minute clock and no coaching.

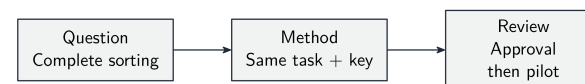
The key is fixed before collection. All four placements must be correct for one complete success. Any wrong or blank placement means no complete success. Teacher approval is pending. Participants may stop or withhold their sheet, no names are collected, and individual sheets cannot be retrieved reliably after anonymous submission.

■ Source B: draft and current evidence

The draft instead asks, “Do you like our new layout?” It places confident readers in B and other readers in A. It promises that every volunteered sheet can be located and deleted later despite collecting no identifier.

Two example sheets are teaching examples, not collected participant data: X has three correct placements and one wrong; Y has four correct placements. A proposed two-person pilot would check wording and timing after approval. No pilot or main study has happened.

■ Worked example: fix the mismatch before a pilot



Known: the question concerns complete sorting success, whereas the draft asks liking. Why use the task? A preference answer cannot supply the required outcome. Under the fixed rule, example X has no complete success and Y has one; do not mix partial marks into the primary count.

Restore the same task and shuffled assignment, explain the anonymous-data limit truthfully, and seek approval. Then pilot instructions and timing. Keep changed-instruction pilot sheets separate from any later main task; do not use them to claim a layout already wins.

2 Practice

2.1 State the outcome measured by the question and the different outcome measured by the draft survey. [2]

2.2 Score teaching sheets X and Y using the complete-success rule. [2]

2.3 Identify THREE relevant controls in the intended task. [3]

2.4 Explain the allocation problem, correct the withdrawal promise and state what a pilot can check. [3]

3 Exam-style questions

3.1 Evaluate THREE claims: “Shuffled slips make the sample represent every student”; “A preference score answers complete sorting accuracy”; “A pilot after changed instructions can be silently merged into the main result”. Give a correction and reason for each. [6]

3.2 Write an 80–110 word method correction. Align question, task and scoring; repair assignment and anonymous-data wording; describe permission and pilot status; preserve the sample limit. [8]

Solutions

2.1 Completely correct four-item sorting; liking/preference.

2.2 X: zero complete successes. Y: one complete success. Three partial correct placements are not three complete successes.

2.3 Award B1 each for three distinct controls: same wording, same item choices/key, same two-minute timing, no coaching, one layout per person. Assignment is a separate allocation procedure; do not repeat timing twice.

2.4 Allocating by reading confidence introduces a group difference tied to condition. Explain stopping/withholding before submission and inability to identify a sheet after anonymous submission. Pilot wording comprehension or workable timing after approval.

3.1 Award 2 per claim: correction, reason. Assignment within class volunteers is not school-wide random sampling. Measure the sorting task: liking is a different outcome. Keep the pilot separate and document changes: different instructions/procedural rehearsal are not the same main comparison.

3.2 Local analytic rubric, $8 \times B1$: accuracy/preference distinction; same task/key; complete-success rule; repaired assignment; truthful anonymous-data limit; approval and separate pilot status; local-sample limit; coherent truthful 80–110 word response. Example: “Our question concerns completely correct sorting, so a liking survey cannot supply the primary result. We will use the same four-item task and fixed key for both layouts, with a two-minute limit and no coaching. A sheet counts as one success only when all four placements are correct. We will use shuffled assignment rather than place confident readers in B. Participants may stop or withhold a sheet; after anonymous submission we cannot reliably retrieve an individual sheet. Teacher approval is still pending. The later pilot will check wording and timing, separately from main data. Class volunteers will not represent every student.”