

Research methods

AS

Read online at pastopic.com/cambridge/psychology-9990/notes/research-methods

Last checked 6 October 2026

© 2026 Pastopic. Free for personal study. Please share the link, not the file.

What you need to know

This is the toolkit for the whole course. Paper 2 is built around it, and the planning question at the end of Paper 4 tests it again. The syllabus asks you to describe each method, evaluate it, and use it in a scenario you haven't seen before ("apply"). By the end of this topic you should be able to:

1. **Experiments.** Describe **laboratory** and **field experiments** and judge each for reliability, validity and ethics. Describe and evaluate the three **experimental designs (independent measures, repeated measures, matched pairs)** and the ideas that go with them: **counterbalancing, random allocation** and **order effects** (practice and fatigue). Explain why experiments use **experimental and control groups** or control conditions.
2. **Self-reports.** Describe **questionnaires** (technique: paper and pencil or online; question format: **open** or **closed**) and **interviews** (format: **structured, semi-structured** or **unstructured**; technique: **face-to-face** or **telephone**; question format: open or closed), and evaluate them.
3. **Case studies.** Describe the method (one participant or one unit, studied in depth) and evaluate it.
4. **Observations.** Describe the choices every observation makes (**overt / covert, participant / non-participant, structured / unstructured, naturalistic / controlled**) and evaluate them.
5. **Correlations.** Describe **positive** and **negative** correlations and their **strength**, name the two **co-variables** and say exactly how each is measured, and evaluate the method, including why it can't show **cause and effect**.
6. **Longitudinal studies.** Describe and evaluate studies that follow the same people over time, including experiments with a longitudinal design.
7. **Plan a study** with any of these methods for a new scenario (Paper 2 Section B, and Paper 4 Section B), giving the features the syllabus lists for that method, and evaluate parts of your own plan.
8. **A Level additions** (Paper 4): **randomised control trials, postal questionnaires, rating scales** and **forced / fixed choice** questions, and **psychometric tests**.

The planning questions also use the other research-methodology topics: aims and hypotheses, variables and controls, sampling, ethics, types of data and data analysis, and validity and reliability. Each of those has its own topic. This note gives you the parts of them you need to plan and to answer method questions, and the "How to do it" section shows how they are asked.

Natural experiments were removed in the 2024–2026 syllabus, so you won't be asked to plan one. They still appear in some 2021–2023 past papers.

Understand it

Why psychologists need different methods

Every method is a trade-off. If you want to know whether one thing **causes** another, you need an experiment, because only an experiment lets you change one thing and keep the rest the same. If you can't (or mustn't) change it, for example a person's age or how much they sleep, you measure what's already there with a correlation. If you want to know what people think or feel, you ask them (a self-report). If you want to see what they actually do, you watch (an observation). If the person is unusual and there's only one of them, you study them in depth (a case study). If you want to see how people change, you follow them over time (a longitudinal study).

A useful way to evaluate any method is to ask four questions about it:

- **Control:** could anything other than what I'm studying have changed the results?
- **Validity:** am I really measuring what I claim to measure, and would people act like this in real life (**ecological validity**)?
- **Reliability:** if I (or another researcher) did it again in the same way, would I get the same results?
- **Ethics:** is anyone harmed, misled, watched without agreeing, or unable to leave?

Experiments

In an experiment the researcher **changes** one thing, the **independent variable (IV)**, and **measures** its effect on another, the **dependent variable (DV)**. Everything else is kept the same with **controls**, so that if the DV differs between conditions, the IV is the likely reason.

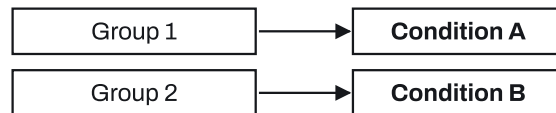
	Laboratory experiment	Field experiment
Where	A setting the researcher has set up, away from where the behaviour normally happens (a lab, a quiet room)	The place where the behaviour normally happens (a shop, a street, a school, a train)
IV	Changed by the researcher	Changed by the researcher
Control	High: the room, the instructions and the timing can all be fixed	Lower: weather, crowds and noise get in the way
Strengths	Easy to repeat exactly, so reliable ; strong evidence that the IV caused the change in the DV	Behaviour is more natural, so higher ecological validity ; people often don't know they are in a study, so fewer demand characteristics
Weaknesses	Artificial tasks and settings, so behaviour may not match real life; people know they are being studied and may guess the aim (demand characteristics)	Hard to control other variables; hard to repeat exactly; participants often can't give consent or be debriefed

Experimental and control groups. To know whether an IV did anything, you need something to compare with. The **experimental condition** gets the thing being tested (a new therapy, music playing); the **control condition** doesn't (no therapy, silence), or gets the usual version. Without a control, an improvement could just be time passing or people trying harder because they are being studied.

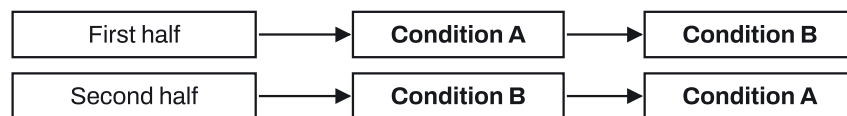
Experimental designs

The design is how participants are shared out between the conditions of the IV.

Independent measures



Repeated measures (counterbalanced)



Matched pairs



- **Independent measures.** Each person takes part in **only one** condition, so you compare separate groups. Strengths: no order effects; people see only one condition, so they are less likely to work out the aim. Weakness: the groups may differ in ways that matter (one may happen to have more able, tired or motivated people). These are **participant variables**. **Random allocation** reduces the problem: deciding by chance (drawing names, a random number generator) which condition each person joins, so that differences between people are spread across the groups by luck, not by the researcher's choice.
- **Repeated measures.** **Every** person does **every** condition, and each person's scores are compared with their own. Strengths: participant variables cancel out, because the same people are in both conditions; fewer participants are needed. Weaknesses: **order effects**. People may do better in the second condition because they have had a go already (**practice effect**), or worse because they are tired or bored (**fatigue effect**). Seeing both conditions also makes the aim easier to guess.
- **Counterbalancing** fixes order effects in repeated measures. Split the participants into two halves: one half does condition A first, then B; the other half does B first, then A (an ABBA pattern). Any practice or fatigue now affects each condition equally, so it can't make one condition look better.
- **Matched pairs.** Before the study, people are **paired** on something that could affect the DV (age, ability, a pre-test score). One member of each pair goes into each condition. Strengths: no order effects (each person does one condition) and participant variables are reduced. Weaknesses: matching takes time; you can never match on everything; if one person drops out, their partner's data is lost too.

Self-reports: questionnaires and interviews

A **self-report** collects data by asking people about themselves.

Questionnaires are written questions that the participant answers alone. The **technique** is how it reaches them:

- **paper and pencil**: printed, filled in while the researcher is around;
- **online**: on a website or app, filled in wherever and whenever the participant likes;
- **postal** (A Level): sent by post and returned by post.

Interviews are questions asked and answered in real time. The **technique** is **face-to-face** or by **telephone**. The **format** is how fixed the questions are:

- **structured**: a set list, asked to everyone with the same wording and in the same order;
- **semi-structured**: a set list of core questions, plus extra questions the interviewer adds to follow up an answer;
- **unstructured**: no fixed list; the interviewer starts with a topic and lets each answer lead to the next question, like a guided conversation.

Format is about **how fixed** the interview is. Whether the questions are open or closed is a separate decision: any of the three formats can use either kind.

Question format applies to both:

	Open questions	Closed questions
What the participant does	Answers in their own words	Picks from answers the researcher has fixed (yes / no, tick boxes, a rating scale)
Example	"What goes through your mind just before an exam starts?"	"Do you revise with music on? Yes / No"
Data	Mostly qualitative (words, detail, reasons)	Quantitative (numbers that can be counted, averaged and compared)
Strength	Rich detail; people can explain <i>why</i> ; nothing is forced on them	Same answer options for everyone, so quick to score, compare and analyse; many people can be asked quickly
Weakness	Hard and slow to analyse; the researcher must interpret the words, which can be subjective	Answers may not fit the person's real view; no reasons are given

Evaluating self-reports. They are the only direct way to find out thoughts, feelings and memories, and a lot of data can be gathered cheaply. But people may give the answer that makes them look good (**social desirability**), may guess what the researcher wants (**demand characteristics**), may misremember, or may not understand a question the way the researcher meant it. Interviews allow questions to be explained, but the

interviewer's manner can influence answers. Online and postal questionnaires reach more people, but there is no one there to clear up confusion, many are never finished or returned, and online ones only reach people with internet access.

A Level extras.

- **Rating scales** turn an opinion into a number, for example "How confident are you? 1 (not at all) to 5 (completely)". A **Likert scale** is a kind of rating scale that measures **agreement** with a statement (strongly agree to strongly disagree). A scale with an even number of points has no middle option, which pushes people to lean one way; a scale with an odd number lets them sit on the fence.
- **Forced (fixed) choice** questions make the person pick from the options given, even if none fits exactly. They are easy to score but may misrepresent what the person really thinks.
- **Psychometric tests** are published, **standardised** tests for measuring a psychological characteristic (such as depression, personality or intelligence), given and scored the same way for everyone and usually checked for reliability and validity. Give the full name before any abbreviation, for example the "Beck Depression Inventory (BDI)".

Case studies

A **case study** investigates **one** participant, or one unit such as a family, a classroom or a small team, in great **depth**, usually over a period of time. The person is often unusual (a rare ability, a rare illness), which is why a study of many people isn't possible.

Case studies normally use **several techniques**: interviews with the person, interviews with people who know them (family, teachers, colleagues), observations, tests, and records such as school or medical files. Using more than one source lets the researcher check whether the sources agree. This is called **triangulation**: if a man's own account of his sleep, his partner's account and a sleep-tracking watch all point the same way, the conclusion is more trustworthy.

- **Strengths**: very detailed (often qualitative) data; can study rare cases that couldn't be studied any other way; triangulation improves validity; can follow change over time.
- **Weaknesses**: one person can't represent everyone, so findings are hard to **generalise**; the researcher may become close to the participant and interpret data subjectively; hard to repeat; unusual people are easy to identify, so **confidentiality** is a constant worry, and talking about painful experiences risks psychological harm.

Observations

In an **observation** the researcher watches and records behaviour. Each observation makes four separate choices:

Choice	One side	The other side
Does the person know they are being watched?	Overt: people can tell they are being watched, and by whom	Covert: the observer is hidden or disguised, so people don't know they are being observed
Does the observer join in?	Participant: the observer takes part in the group's activity	Non-participant: the observer watches from outside the activity
How is behaviour recorded?	Structured: a fixed list of behavioural categories , each counted (tallied) when it happens	Unstructured: the observer writes down anything that seems relevant
Where does it happen?	Naturalistic: the normal setting, which the researcher doesn't change	Controlled: a setting the researcher has set up or arranged

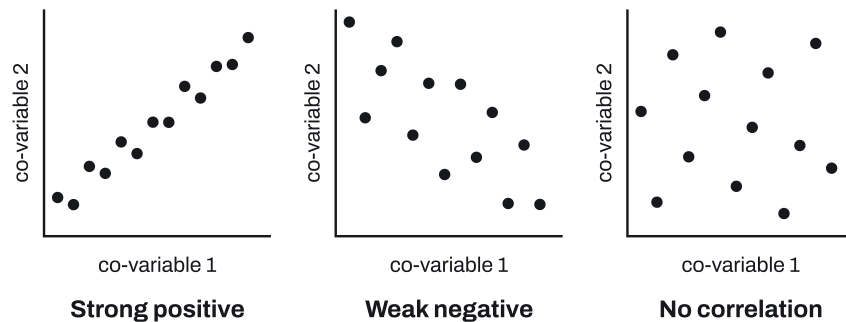
The four choices are independent. An observer can be **non-participant and overt** (a researcher sitting openly at the back of a lesson) or **participant and covert** (a researcher who has taken a job in a café without telling anyone). Non-participant does **not** mean covert.

Behavioural categories are what make a structured observation work. Each one must be a single behaviour that can be **seen** and that two observers would recognise in the same way: "laughs out loud" is a good category, "is happy" is not, because happiness can't be seen directly. Clear categories improve **inter-observer reliability**, which is checked by having **two or more observers** record the same behaviour separately and comparing their tallies. A plan can also say how behaviour is sampled: record every time a category happens (**event sampling**) or record what is happening at fixed moments, such as every 30 seconds (**time sampling**).

- **Overt vs covert:** overt is more ethical (people can agree and can withdraw), but people may change how they act once they know they are watched. Covert behaviour is more natural, but people can't consent or withdraw, and watching them may invade their privacy.
- **Participant vs non-participant:** joining in gives insider understanding, but the observer may lose objectivity or find it hard to record in the moment. Watching from outside is more objective but may miss what behaviour means to the group.
- **Structured vs unstructured:** categories give quantitative, comparable and reliable data but may miss behaviour that isn't on the list. Unstructured observation catches everything but is subjective and hard to compare.
- **Naturalistic vs controlled:** natural settings give higher ecological validity; controlled settings make it easier to see, record and repeat.

Correlations

A **correlation** looks for a **relationship** between two variables that are both **measured**, not changed. They are called **co-variables**. Each participant gives a score on both, and the pairs of scores are plotted on a **scatter graph**.



Schematic: made-up points to show the patterns

- **Positive correlation:** as one co-variable increases, the other **increases** too (more practice, higher score).
- **Negative correlation:** as one co-variable increases, the other **decreases** (more days absent from class, lower end-of-term test score).
- **Strength:** the closer the points lie to a straight line, the stronger the correlation. Widely scattered points mean a weak correlation; no pattern means no correlation.

Both co-variables must be measured on a **scale** (a count, a time, a rating, a test score), not as yes / no categories, or there is nothing to plot. Each one needs an **operational definition**, an exact statement of how it will be measured: not "fitness" but "the number of minutes the person can jog without stopping".

Correlation does not show cause. Even a strong correlation can't tell you which variable affects which, or whether a third variable is behind both. If time spent reading and vocabulary size go up together, reading might build vocabulary, or a big vocabulary might make reading more enjoyable, or both might come from growing up with lots of books at home.

- **Strengths:** can study variables that can't or mustn't be changed (for ethical or practical reasons); shows whether a relationship is worth testing with an experiment; uses data that is often easy to collect.
- **Weaknesses:** no cause-and-effect conclusion; a relationship that isn't a straight line may be missed; the measures (often self-reports) may not be valid.

Longitudinal studies

A **longitudinal study** tests the **same participants** again and again over a **long period** (months or years), to see how they change. An **experiment with a longitudinal design** does the same with an IV, for example a new therapy, measuring the DV at regular points during treatment.

When you plan one, decide: the **tests or tasks** used at each point and how they are **scored**; how **often** people are retested (the **interval**: often enough to catch the change, not so often that people get bored or learn the test); how participants will be **recontacted** (and what could stop that); and the **controls** that keep each testing session the same.

- **Strengths:** shows change within the same people, so participant variables stay the same across the study; can show the order in which changes happen.

- **Weaknesses: attrition** (people move away, change their phone number or email, become ill, or lose interest, and those who stay may differ from those who leave); slow and expensive; repeated testing can cause practice effects; the world may change during the study.

Comparing **different age groups at the same moment** (a cross-sectional comparison) is quicker and has no drop-out, but the groups differ in more than age.

Randomised control trials (A Level)

A **randomised control trial (RCT)** is the standard way to test whether a treatment works. Participants are **randomly allocated** to the treatment group or to a **control** group (no treatment, a placebo, or the usual treatment), and the outcome is measured in both. It is often **single-blind** (participants don't know which group they are in) or **double-blind** (neither the participants nor the researchers who measure the outcome know). Random allocation spreads participant variables evenly, and blinding stops expectations shaping the results. The main concerns are ethical: the control group goes without a treatment that might help them.

Planning a study

The planning question gives you a method and a topic, and you write a plan someone else could follow. The syllabus splits what to include into two kinds of feature.

Specific features depend on the method:

Method	Features to plan
Experiment	Type of experiment; IV (each condition described); DV (exactly how it is measured); controls / standardisation; experimental design; counterbalancing or random allocation where they fit; for a field experiment, the location
Questionnaire	Technique (paper and pencil, online; postal at A Level); question format (open, closed); example questions; how answers are scored or interpreted
Interview	Format (structured, semi-structured, unstructured); technique (face-to-face, telephone); question format; example questions; scoring or interpretation
Case study	Who the participant or unit is; what information is collected; two or more data-collection techniques; how the data is analysed, interpreted and triangulated
Observation	Overt / covert; participant / non-participant; naturalistic / controlled; structured (with the behavioural categories) / unstructured; at A Level, the number of observers
Correlation	The two co-variables; how each is measured; what will show the relationship (scatter graph)
Longitudinal study	Tests or tasks; scoring; frequency / interval; recontacting participants; controls / standardisation

General features apply to any method: the **procedure**; how the plan makes the study **valid** and **reliable**; **types of data**. Paper 4 adds the **aim and hypothesis**, the **sample and sampling technique**, **ethics**, and **analysis with descriptive statistics**. Paper 2 doesn't expect an aim or hypothesis in the plan, or an evaluation of its sample or ethics, and its questions usually tell you not to describe the sample or ethics.

Key facts and formulas

There is no formula list in the 9990 exam. "Learn it" means you must remember it.

Fact	Status
Laboratory experiment: IV changed by the researcher in a setting they have set up; high control, lower ecological validity	Learn it
Field experiment: IV changed by the researcher in the place the behaviour normally happens; higher ecological validity, less control	Learn it
Independent measures: each participant does one condition. Repeated measures: each participant does every condition. Matched pairs: participants paired on a relevant variable, one of each pair per condition	Learn it
Counterbalancing with n participants: $\frac{n}{2}$ do condition A then B, and $\frac{n}{2}$ do B then A, to balance order effects	Learn it
Order effects: practice (better second time) and fatigue (worse later, through tiredness or boredom)	Learn it
Questionnaire technique: paper and pencil, online (postal at A Level). Interview technique: face-to-face, telephone	Learn it
Interview format: structured (same questions, same order), semi-structured (fixed core plus follow-ups), unstructured (no fixed list)	Learn it
Open questions: own words, mostly qualitative data. Closed questions: fixed options, quantitative data	Learn it
Case study: one person or unit, in depth, several techniques, triangulation	Learn it
Observation choices: overt / covert, participant / non-participant, structured / unstructured, naturalistic / controlled	Learn it
Positive correlation: both co-variables rise together. Negative: one rises as the other falls. Strength: how closely points follow a straight line. No cause and effect	Learn it
Longitudinal study: same participants, tested repeatedly, over a long period; main problem attrition	Learn it
RCT: random allocation to a treatment or control group, often single- or double-blind	Learn it

Fact	Status
Range = highest score – lowest score	Learn it

How to do it

We have tagged 307 questions on research methods, from 2021 to 2026: 167 on Paper 2 and 140 on Paper 4 (the Paper 4 planning questions are printed in more than one paper variant, so 40 of the 140 are repeats). 238 of them have parts of more than one type, and every Paper 4 planning question counts towards the first three types, so the counts add up to 709.

Question type	Questions
Plan a study for an applied scenario (Paper 4 (a))	140
Describe the psychological knowledge behind your plan (Paper 4 (b))	140
Justify a choice in your own plan (Paper 4 (c))	140
Identify and operationalise variables	30
Self-reports: questions, techniques and formats	26
Ethics in a scenario	26
Describe how to conduct a study (Paper 2, 10 marks)	25
Evaluate your own procedure (Paper 2)	25
Describe two methods with examples (6 marks)	19
Data: types, tables, graphs and statistics	19
Experimental designs	18
Observations	17
Validity and reliability	16
Controls and uncontrolled variables	15
Correlations	14
Case studies	11
Types of experiment	9
Sampling	8
Hypotheses	7
Longitudinal studies	4

Find your question type by area:

- **Planning and evaluating your own study:** types 1, 2, 3 (Paper 4), 7 and 8 (Paper 2).
- **The methods themselves:** 5 (self-reports), 9 (describe two methods), 12 (observations), 15 (correlations), 16 (case studies), 17 (types of experiment), 20 (longitudinal studies).
- **Experiment design and variables:** 4 (variables), 11 (designs), 14 (controls), 19 (hypotheses).
- **Quality of a study:** 6 (ethics), 13 (validity and reliability), 18 (sampling).
- **Data:** 10.

1. Plan a study for an applied scenario (140 questions)

Paper 4 Section B: you answer **one** of four questions, one per option. Since 2024 it reads "Plan a study using *method* to investigate *topic*. Your plan must include details about *feature 1* and *feature 2*" (10 marks for the plan); from 2021 to 2023 it was "Design a study..." with no named features. The plan is marked on how well you apply method features, both the specific features of the named method and general features.

1. **Use the method named in the question.** An interview means an interview, not an experiment with an interview tacked on. Some candidates turn every plan into an experiment, and lose marks for it.
2. **Cover both named features in detail**, because part (c) will ask you to justify them.
3. **Give four or five specific features** of that method (see the table in "Planning a study"), each with a reason or detail, not a bare list. "The design is independent measures" on its own earns little; add how people are put into each group and why that design suits the question.
4. **Add the general features:** an aim and a hypothesis, the sample (who, how many, from where) and the sampling technique, the location, the procedure step by step, ethics (consent, right to withdraw, debrief, confidentiality), the type of data and how you would summarise it (mean, bar chart), and one or two steps that make it valid and reliable (standardised instructions, two observers).
5. **Tie it to the topic.** Your questions, categories or DV must actually measure what the question asks about (for example, how effective a treatment is, not just what the symptoms are).
6. **Don't evaluate and don't invent results** in (a). Evaluation is what (c) is for, and a plan has no results yet.

Short worked plans for each method:

- **Experiment:** type (lab or field), IV with each condition, DV with its unit, two or three controls, design, then counterbalancing (repeated measures) or random allocation (independent measures), and single- or double-blind if it is a treatment trial.
- **Questionnaire:** technique, number of questions, question format with two real example questions written out, how answers are scored (for example 1–5 per item, added up), how the score is interpreted.
- **Interview:** format, technique, question format with examples, who asks, where, how long, how answers are recorded and analysed (scores for closed questions; open answers sorted into themes by two independent raters).
- **Observation:** overt or covert, participant or not, naturalistic or controlled, structured with three or four named behavioural categories, event or time sampling, two observers who compare tallies.

2. Describe the psychological knowledge behind your plan (140 questions)

Part (b) links the plan to the option topic you studied.

1. **(b)(i), 4 marks:** describe the **psychological** knowledge from the syllabus that the question is about: the theory, the treatment, the key study or the named example study. Short, thin answers stay in the bottom band (1–2 marks); you need several accurate, detailed points for 3–4. Describing methodology here earns nothing, and so does a theory from the wrong part of the syllabus.
2. **(b)(ii), 4 marks:** pick **two features** of what you wrote in (b)(i) and, for each, say where it shows up in your plan. One mark for the feature, one for showing how it shaped (a). For example: "In (b)(i) I described the two kinds of treatment, so in (a) my IV had exactly those two conditions."
3. Anything you use in (b)(ii) must already be in (b)(i).

From 2021 to 2023 this was one 8-mark part asking for the "psychological and methodological evidence" behind the plan: explain the theory or study it rests on, then explain two or three design decisions, each with a reason, rather than listing them.

The knowledge itself belongs to the option topics (clinical, consumer, health, organisational), not to this note.

3. Justify a choice in your own plan (140 questions)

Part (c) is three 2-mark parts about the features named in the question: "one reason for your choice of X", "one weakness of your choice of X", sometimes "one strength". The rule is the same each time:

1. Use **exactly the choice you made in (a)**. If (a) used closed questions, (c) must be about closed questions.
2. Give a reason, strength or weakness that is true of that choice in general (**1 mark**).
3. Link it to **your** study, using its details: the participants, the topic, an example question (**2 marks**).

Ready-made points (always add the link to your plan):

Choice	A reason or strength	A weakness
Closed questions	Every answer fits the same fixed options, so scores can be added up and compared quickly	People can't explain their answer or give a view the options don't include
Open questions	Participants can give reasons and detail in their own words	Slow to analyse; the researcher has to interpret the words
Structured interview	Standardised: the script and its order never change between participants, so interviews can be compared and repeated	Can't follow up an interesting answer
Semi-structured interview	Fixed core questions plus the freedom to ask follow-ups	Different follow-ups make interviews harder to compare

Choice	A reason or strength	A weakness
Unstructured interview	Total flexibility to explore what each person says	Hard to compare and analyse; depends on the interviewer
Online questionnaire	Easy to send to hundreds of people in many places, and nobody stands over them while they answer	Leaves out anyone without a device or internet connection, and it is easy to close the page before finishing
Postal questionnaire	Reaches people spread over a wide area, who answer at home in their own time	Many are never sent back: they get put aside and forgotten, or returning them costs the participant money
Paper and pencil	The researcher hands it out and waits, so fewer go missing and confusing questions can be explained	With the researcher in the room, people may tailor answers to what they think is wanted
Random sampling	Each member of the target population is as likely to be picked as any other, so the researcher's preferences play no part	You need a list of the whole population, and those picked may refuse
Opportunity sampling	Quick and easy to get enough people	Whoever the researcher decides to approach (for example, only people who seem friendly) can skew the sample
Volunteer sampling	Those who reply have already chosen to take part, so fewer drop out or refuse	People who put themselves forward may share traits, such as being keen to help, that make them unlike the rest
Laboratory experiment	Other variables are fixed, so a difference between conditions can be put down to the IV with more confidence	Artificial, so low ecological validity
Field experiment	The setting is the real one, so behaviour is closer to everyday life	Harder to control other variables
Independent measures	No one does both conditions, so there are no order effects; suits a treatment-versus-control comparison	Participant variables differ between groups
Repeated measures	Controls participant variables	Order effects (needs counterbalancing)
Controls	Remove other explanations for a change in the DV	Every extra control makes the situation less like real life
Directional hypothesis	Earlier research suggests which way the difference will go	—
Non-directional hypothesis	Past findings are missing or point both ways, so you can't predict which group will come out ahead	—

Choice	A reason or strength	A weakness
Behavioural categories	Settled in advance, so two observers tally the same actions and their records can be checked against each other (inter-observer reliability)	Behaviour outside the list is missed
Quantitative data	Can be summarised with descriptive statistics and compared	Says nothing about reasons
Lab or hospital room as location	Noise, interruptions and equipment can be kept the same for everyone	An unfamiliar room may make people tense, so they act differently from at home

Variations: a hypothesis justification earns the marks only if the hypothesis itself is directional or non-directional and fits the plan; a **null hypothesis earns 0**. "Sample of participants" (how many, ages, where from, what kind of people) is **not** the same as "sampling technique" (random, opportunity, volunteer): answer the one asked.

4. Identify and operationalise variables (30 questions)

1. **Identify the IV:** what the researcher changes, or the thing that differs between the conditions. Name **both** conditions ("music on vs no music"), not just the topic.
2. **Identify the DV:** what is measured, with how it is measured.
3. **Operationalise** = turn a vague idea into something you can measure or set exactly. "Stress" becomes "heart rate in beats per minute after the task", or "score on a 1–10 rating of how tense the person feels". An operationalised IV gives each condition's exact value ("a 10-minute walk vs 10 minutes sitting").
4. **Strength or weakness of a measure:** objective measures (a count, a time, a machine reading) can't be biased by the researcher's opinion; self-rated measures are easy to collect but open to social desirability, and a single number may not capture the whole idea.

"Feeling relaxed" as quantitative data: ask for a rating on a numbered scale. "Physical activity": steps per day from a phone app.

5. Self-reports: questions, techniques and formats (26 questions)

1. **Open or closed?** Any question with fixed answers (yes / no, tick boxes, a rating scale) is **closed**. A question that starts "Describe...", "Explain..." or "Why..." and can't be answered in a word is **open**.
2. **Write one of each** when asked, on the topic given. If told "do not use a rating scale", make the closed question yes / no or multiple choice.
3. **Questionnaire or interview:** questionnaires are cheaper and reach more people and the researcher can't influence the answers face to face; interviews let you explain questions and follow up answers.
4. **Questionnaire technique (AS):** the two are **paper and pencil** and **online**. A reason for paper: it feels more private and the researcher can collect it; for online: easy to reach a wide sample, nothing to post back.

5. **Interview formats:** say what makes an interview structured (same questions, same order, same wording for everyone) and how to turn an unstructured one into a structured one (write a fixed list and ask it in order).

6. Ethics in a scenario (26 questions)

Each point is usually **1 mark for the guideline + 1 mark for the link** to the scenario.

1. Name a guideline: **minimising harm, valid (informed) consent, right to withdraw, lack of deception, confidentiality, privacy, debriefing.**
2. Say how it applies in **this** study, using the scenario's details.

Watch which guidelines actually fit. In a covert observation in a public street, asking for **consent** would alert people and ruin the natural behaviour you want to record, so it usually isn't sought; **debriefing** may be impossible once people walk off; and someone who notices being watched may feel uneasy (**harm**).

Deception needs someone to be misled, so it rarely fits when nobody speaks to the observer. **Privacy** is seldom broken by watching people on a public street, since anyone there can be seen by passers-by; it would matter if the observer wrote down something personal, such as two friends' conversation. Using an initial or a code instead of a name ("Mr K") is about **confidentiality**. Recalling painful memories brings in **minimising harm** and the **right to withdraw**. With children, ask the **parent's** consent and also the **child's** agreement in words they understand, and treat signs of upset as wanting to stop.

Animals: the guidelines are minimising harm, **replacement** (use something other than live animals where possible, such as videos of animals in the wild), **species** (choose animals likely to suffer least, such as ones with simpler nervous systems), **numbers** (use as few as possible) and **procedures** (keep pain, suffering and distress low; suitable housing; care with rewards, deprivation and unpleasant stimuli).

7. Describe how to conduct a study (25 questions)

Paper 2 Section B (a), 10 marks: "Describe how *a researcher* could conduct a *method* to investigate...", usually adding "do not describe the sample, sampling technique or ethics". It is marked in levels against the method's **required features** (the specific features in the "Planning a study" table): Level 5 (9–10 marks) needs **every** required feature, **each with detail**, using correct terms.

1. List the required features for the method before you write. For a correlation: two co-variables; measure of variable 1; measure of variable 2; how the relationship is shown.
2. Write each one with real detail: exactly what is measured, with what, scored how.
3. Keep it replicable: someone else should be able to run your study from your description.
4. Leave out what the question excludes.

For a **correlation**, don't call the variables IV and DV and don't split people into groups (that's an experiment). Both measures must give **numbers** on a scale; open questions or an unstructured observation don't work.

For a **case study**: who the person is (age, how they were found), what you want to know about them, at least two techniques (interview them, interview family or colleagues, observe them, test them), and how the data is combined and interpreted (triangulation).

Other methods have been asked: a field experiment, an independent-measures experiment, a structured observation, a structured interview, and (before 2024) a natural experiment.

8. Evaluate your own procedure (25 questions)

Paper 2 Section B (b), about the plan you just wrote. Ethics and sampling are excluded.

- **2025–2026 style:** a strength and a weakness of **one part** of your procedure (both parts about the **same** part), or two features that make it valid. Each is 1 mark for the point + 1 mark for its effect on the results. Good parts to choose: how a variable was operationalised, a control, the design, a standardised procedure, triangulation.
- **Two strengths of your study:** pick two different features of your procedure (for example a standardised script and an objective measure of the DV) and, for each, say how it improves the results.
- **Independent vs repeated measures in this study:** give one advantage and one disadvantage of using independent measures rather than repeated measures, each tied to your study. For example: no order effects, because each person does only one condition; but participant variables, such as differences in ability between the two groups, could affect the DV.
- **2021–2023 style (4 marks):** one practical weakness of your procedure and how you would change it to fix the weakness.

Don't praise a correlation for producing quantitative data: every correlation needs it, so it's not a strength of *your* plan.

9. Describe two methods with examples (19 questions)

A 6-mark "Describe X and Y, using any example(s)" question. Up to 2 marks for describing each term (a definition plus a point of detail), and up to 2 for examples of each. With no examples, or only one term covered, the most you can get is 4.

1. **X:** a definition, then one more detail (a strength, the data it gives, how it works).
2. **An example of X:** from a core study, another study or one you invent ("asking 'Describe how you felt when...'").
3. Do the same for **Y**.

Pairs asked so far include: open and closed questions; the three interview formats in pairs; paper-and-pencil and online questionnaires; laboratory and field experiments; matched pairs and repeated measures; participant and non-participant, covert and overt, and structured and unstructured observations; positive and negative correlations; correlations and experiments.

10. Data: types, tables, graphs and statistics (19 questions)

- **Quantitative** data is numbers; **qualitative** data is words or descriptions. Give an example of each from the scenario.
- **Draw a table for an observation:** one column (or row) per behavioural category, with a row for the **tally** or **frequency** of each. A table without a way of counting, or without the categories, loses marks.
- **Graphs:** a **bar chart** for separate categories (bars with gaps), a **histogram** for continuous data, a **scatter graph** for a correlation. Label both axes with the variable (and units), and plot the values exactly.
- **Mode** = most frequent score; **median** = the middle score when they are put in order (with two middle numbers, take halfway between them); **mean** = the arithmetic average. **Range** = highest minus lowest. **Standard deviation:** the larger it is, the more spread out the scores are around the mean.
- A **conclusion** says what the data **shows** about behaviour ("people helped more when..."), not the numbers themselves, which are a result.

11. Experimental designs (18 questions)

1. **Define** the design and give the scenario's example ("each child played with only one of the two toys").
2. **Identify** it from a description: same people in both conditions = repeated measures; different people = independent measures; different people paired up first = matched pairs.
3. **Strength or weakness**, or why one design beats another, with a link (see "Experimental designs" above).
4. **Counterbalancing:** define it (half do the conditions in one order, half in the reverse order) and say **why it's needed here:** which order effect it removes in this scenario. Describing how to do it is not the same as explaining why.
5. **Random allocation:** name it and say what it achieves (participant variables spread evenly by chance).
6. **Matching:** name variables that would affect the DV in this scenario.

12. Observations (17 questions)

1. **Classify** the observation on each dimension asked, with a reason from the scenario ("covert, because the shoppers didn't know the observer was there").
2. **Strength or weakness** of that choice, compared with the alternative when the question says "better than".
3. **Suggest behavioural categories** that are single, visible actions ("claps", "checks a phone"), not broad ideas ("body language") or feelings.
4. A **practical problem** with a category: it may be hard to see from a distance, or it could be confused with something else.

"Overt" means people know an observer is there; they may still not know the aim. People in an overt observation don't "show" demand characteristics, they **respond to** them.

13. Validity and reliability (16 questions)

- **Validity:** does the measure capture what it claims to? Don't stop at the definition. Look at the scenario: a passer-by may be smiling at a message on their phone, not at the person they pass, so counting smiles could overstate friendly greetings; a questionnaire about kindness may get socially desirable answers rather than real behaviour.
- **Reliability:** is the measure consistent? A problem with **reliability** is about people (or one person on two occasions) recording or scoring the same thing differently; a problem with **validity** is about measuring the wrong thing.
- **Objectivity:** recording facts not opinions (a timer, a count), so different researchers would agree.
- **Standardisation:** everyone gets the same instructions, questions, timing and setting. For an interview: a script, the same interviewer, the same order of questions.
- **Improving validity:** use more than one technique (triangulation); keep participants unaware of their condition so their expectations can't affect the DV; check with a second observer.

14. Controls and uncontrolled variables (15 questions)

1. **Situational variable:** something in the **environment** that could affect the DV (noise, temperature, lighting, time of day). Name the specific thing ("the noise from the drilling"), not just the source.
2. **Participant variable:** something about the **people** that could affect the DV (age, ability, mood, experience).
3. **Uncontrolled variable:** any variable like these that the study didn't keep the same.
4. **How to control it:** keep it the same in every condition (same room, same time of day), or use the design (repeated measures, matched pairs, random allocation).
5. **Why it matters here:** explain how it could change this DV in this scenario ("a hotter room might make people irritable, so they would stand further from others").

15. Correlations (14 questions)

- **State the relationship in context:** "a negative correlation means the more hours of exercise, the lower the anxiety score".
- **Why a correlation and not an experiment:** the variable can't be changed (age, personality), or changing it would be unethical (making people lonely).
- **Data must be on a scale:** yes / no answers can't be plotted as a correlation.
- **Strength from a scatter graph:** how close the points are to a straight line.
- **No cause and effect:** either variable could affect the other, or a third variable could affect both.

16. Case studies (11 questions)

1. **Features:** one participant (or unit), studied in depth, usually over time, using several techniques.
2. **Techniques:** interviews (with the person and people who know them), observation, tests, records.

3. **Triangulation:** compare data from different techniques or sources to see whether they agree; give an example using the scenario's sources.
4. **Strength or weakness:** depth and detail vs one person, so no generalising; subjective interpretation; hard to replicate. **Ethical weaknesses:** the person may be identifiable, and probing in depth can cause distress.

17. Types of experiment (9 questions)

1. **Say why** a study is a lab or field experiment: was the IV changed by the researcher, and was the setting artificial or the normal one for that behaviour?
2. **Strengths and problems:** in the field, the DV is harder to measure and other variables are harder to control, but behaviour is more natural; in the lab, the reverse.
3. Before 2024 you could also meet **natural experiments**, where the IV varies naturally and the researcher doesn't change it. They are no longer in the syllabus.

18. Sampling (8 questions)

- **Population** = the whole group the researcher wants to find out about; **sample** = the people who actually take part.
- **Random sample:** get a list of the whole target population (a school register, a staff list), then pick by chance (random number generator, names from a hat) until you have enough. Describe it in the scenario's terms.
- **Opportunity sample:** whoever is available and willing at the time; quick, but not representative.
- **Volunteer sample:** people who respond to an advert (a poster, an email, a social media post) and get in touch. Say where the advert goes and how people reply.
- **Why recruit from several places:** one place may be unusual, so a wider sample is more **representative** and the results **generalise** better. Link it to what is being measured.
- **Inclusion criteria:** the conditions a person must meet to take part (an age range, a diagnosis), so the sample fits the aim.

19. Hypotheses (7 questions)

- **Directional** (one-tailed): says **which way** the result will go ("people who walk to work will report lower stress scores than people who drive").
- **Non-directional** (two-tailed): says there **will be a difference** (or a relationship), but not which way.
- **Null:** says there will be **no** difference (or no relationship), or that any difference is due to chance. Include both operationalised variables.
- **Choosing:** directional when earlier research points one way; non-directional when it doesn't, or results so far disagree.

20. Longitudinal studies (4 questions)

1. **What makes it longitudinal:** one group of people is followed for months or years and measured again at several points, so the **same participants** are compared with themselves; then point to where the scenario shows this (for example, "the same patients are checked every month").
2. **Retest interval:** suggest a fixed interval (every six months, every year) and justify it: often enough to see the change, not so often that people get tired of it or drop out. Weekly is usually too frequent for slow changes like development; ten years is too rare.
3. **Recontacting:** keep several contact details; problems are moving house, a new phone number or email, illness, or no longer wanting to take part.
4. **Compared with testing different age groups at one time:** quicker, no drop-out, but the groups differ in more than age.

Worked examples

Example 1

Rafael wants to know whether people type a short passage faster on a phone screen or on a laptop keyboard. He tests 24 students, and every student types the passage on both devices.

(a) Name the experimental design Rafael is using and outline what it means, using his study as an example. [2]

(b) Explain why Rafael should use counterbalancing in his study. [2]

(c) Suggest one situational variable that Rafael should control, and how he could control it. [2]

(a) Repeated measures: each participant takes part in every condition of the IV **(1)**. Here every student types the passage on both the phone and the laptop, and their two typing times are compared **(1)**.

(b) The second time a student types the passage they already know it, so they may type it faster whichever device comes second (a practice effect) (1). If 12 students use the phone first and 12 use the laptop first, any practice benefit falls equally on both devices, so it can't make one look faster **(1)**.

(c) Noise in the room (1): test every student in the same quiet room, at the same time of day, with the same printed passage **(1)**.

Example 2

Interviews can be carried out face-to-face or by telephone. Describe each of these interview techniques, giving an example of each. [6]

Face-to-face: the interviewer and participant meet and talk in the same room, asking and answering questions in real time **(1)**. Because they can see each other, the interviewer can notice confusion or upset and adjust, but their presence may make the participant give answers that make them look good **(1)**. Example: a

researcher sits with a nurse in a quiet hospital office and asks how she copes with night shifts **(1)**.

Telephone: the questions are asked and answered over a phone call, so the two people can't see each other **(1)**. This makes it cheaper and quicker to reach people living far away, and some people find it easier to talk honestly without being watched, but non-verbal signs such as facial expressions are lost **(1)**. Example: a researcher phones farmers in remote villages and asks how often they feel lonely during the harvest season **(1)**.

Example 3

Tasha thinks that cyclists waiting at a red traffic light behave more impatiently when they ride alone than when they ride in a group.

(a) Describe how Tasha could carry out a structured observation to investigate this. Do not describe the sample, sampling technique or ethics. **[10]**

(b) For one part of the procedure you described in (a), describe one strength and one weakness of that part. Do not refer to sampling or ethics. **[4]**

(a) A Level 5 answer covers every required feature of an observation, each with detail:

- **Naturalistic, non-participant, covert:** Tasha sits on a bench at a busy city-centre crossing with a cycle lane, looking like someone waiting for a friend, with a notebook, and doesn't ride or speak to anyone. Cyclists don't know they are being observed, so they behave as they normally would. (Naturalistic / controlled, participant / non-participant, overt / covert, each with detail.)
- **Structured, with behavioural categories:** from the moment a cyclist stops at the red light until it turns green, she records whether they are alone or riding with others, and tallies every time one of these categories happens during the red phase (event sampling): "rolls forward past the stop line", "rings bell", "looks up at the light", "stands still with a foot down". "Rolls forward" means the front wheel crosses the white line.
- **Recording:** a table with one row per cyclist and a column per category, plus the length of the red phase measured with a stopwatch.
- **Reliability and analysis:** a second observer on the other side of the road records the same cyclists separately, and the two tallies are compared. Each cyclist's impatient acts (rolling forward plus ringing the bell) are added up, and the mean number of impatient acts per cyclist for lone riders is compared with the mean for group riders on a bar chart.

(b) Part: the behavioural category "rings bell". **Strength:** it is a single action, defined in advance, that is easy to see and hear, so both observers know exactly what to tally **(1)**; this makes their tallies more likely to agree, improving inter-observer reliability **(1)**. **Weakness:** a cyclist may ring the bell to warn a pedestrian stepping into the lane, not out of impatience **(1)**, so some tallies would count the wrong thing and the results may overstate impatience **(1)**.

Example 4

This is Paper 4 style. Part (b) of the real question asks for knowledge from your option topic, so only (a) and (c) are shown.

(a) Plan a study using a postal questionnaire to investigate how satisfied retired people are with the leisure activities offered at community centres. Your plan must include details about: question format; sampling technique. [10]

(c)(i) Explain one reason for your choice of question format. [2] (c)(ii) Explain one weakness of your choice of sampling technique. [2] (c)(iii) Explain one weakness of using a postal questionnaire in your study. [2]

(a) A plan with four or five specific features of a questionnaire, plus general features:

- **Aim and hypothesis:** to find out how satisfied retired people are with community-centre activities. Non-directional hypothesis: there will be a difference in satisfaction scores between people who attend weekly and people who attend less often.
- **Sampling technique and sample:** a random sample. With permission, each of five centres in one city gives its list of registered members aged 65 or over; every name gets a number, and a random number generator picks 40 from each centre, 200 in all.
- **Technique:** a postal questionnaire with a stamped, addressed return envelope, so people can take their time at home.
- **Question format:** closed questions on a 5-point rating scale, for example: "How satisfied are you with the exercise classes? 1 (very dissatisfied) to 5 (very satisfied)", and "How many times a month do you visit the centre? 0 / 1–2 / 3–4 / 5 or more". Ten satisfaction items.
- **Scoring:** the ten ratings are added, giving a satisfaction score from 10 to 50; higher means more satisfied.
- **Ethics:** a cover letter explains the aim, that taking part is optional, that the questionnaire has no name on it (confidentiality), and that it can be left unfinished; a debrief sheet goes with the results.
- **Analysis:** mean and range of the satisfaction scores for weekly and less-frequent visitors, shown on a bar chart.
- **Reliability:** every person gets the same printed questions, in the same order, with the same instructions.

(c)(i) Closed rating-scale questions give every person the same answer options, so the scores are easy to add up and compare (1); this lets me compare the mean satisfaction score of weekly visitors with that of less-frequent visitors directly (1).

(c)(ii) A random sample needs a full list, and people picked at random may not agree to take part (1); members who rarely go to the centre may be the least likely to post the questionnaire back, so the less-frequent group could end up small and unrepresentative (1).

(c)(iii) A large share of posted questionnaires never come back: people mean to fill them in, then forget, or decide it isn't worth the effort (1); retired people who are less satisfied, or who find writing difficult, may be the ones who don't send it back, so the satisfaction scores could look higher than they really are (1).

Common mistakes

- **Using the wrong method.** If the question says interview, questionnaire or observation, don't turn the plan into an experiment. Examiner reports often mention this.
- **Mixing up format and technique.** Interview **format** is structured / semi-structured / unstructured; interview **technique** is face-to-face / telephone; **question format** is open / closed; questionnaire **technique** is paper and pencil / online / postal. Answer the one the question names.
- **Thinking the format decides the question type.** "Structured" is about a fixed script; the questions on it may be open, closed or a mix, and a semi-structured interview can produce purely numerical answers or purely words.
- **Muddling open and closed questions** when writing examples. A question with answer options is closed, even if it looks long. If you stick to a single question format throughout the plan, your example questions and your part (c) answers can't contradict each other.
- **Calling a correlation's variables IV and DV**, or splitting people into high and low groups. That describes an experiment. Use co-variables, both measured on a scale.
- **Repeating the term as its definition.** "Counterbalancing is balancing the conditions" explains nothing. Say what is actually done and what it achieves, with an example.
- **Saying "not valid" or "not reliable" without explaining.** Say why, in this scenario: what could be measured instead of what was intended, or who might record it differently.
- **Generic answers.** When a part names a scenario or your own plan, the second mark usually goes to the link: a correct point that could apply to any study tends to earn only half.
- **Thinking non-participant means covert.** They are separate choices.
- **Misusing ethics.** Choosing volunteers is a sampling decision and doesn't by itself meet any ethical guideline; privacy is rarely a problem in a public place; deception needs someone to be misled.
- **"Sample" vs "sampling technique".** The sample is who and how many; the technique is how they were chosen.
- **Answering two Section B questions** on Paper 4. Only one is marked.
- **Evaluating or giving results inside a plan.** The plan earns nothing for either.
- **Explaining how counterbalancing works when asked why it's needed.** Say which order effect it removes in this study.
- **Using a histogram for categories.** Separate categories need a bar chart with gaps.
- **Giving a result instead of a conclusion.** Quoting the numbers is a result; a conclusion says what they show about behaviour.
- **Changing the part of the procedure** between the strength and the weakness in Paper 2 (b), when the question says "this part".

How to get the marks

- **"Identify" or "name":** a word or phrase is enough ("repeated measures").

- **"Outline" or "define"**: one accurate sentence, in your own words, that says what the term means, not just repeats it.
- **"Describe"**: enough separate points to match the marks. In "Describe X and Y, using any example(s)", give two points and an example for each term.
- **"Explain"**: make the point, then link it, usually to the scenario ("...so in Mia's study..."). The second mark is almost always the link.
- **"Suggest"**: apply what you know to the scenario and give a reason; there's no single right answer, but it must fit the situation.
- **"Plan" (Paper 4) and "describe how ... could conduct" (Paper 2)**: every required feature of the named method, each with detail, in correct terms, in a form someone could replicate.
- **"Evaluate"**: strengths and weaknesses, each explained and linked.
- **Part (c) on Paper 4**: a generic point gets 1 mark; tied to your plan it gets 2. It must match what you wrote in (a).
- **Ethics questions**: guideline (1) + how it applies here (1).
- **Graphs**: the right type, both axes labelled, sensible scale, values plotted exactly; tables need full headings and a way of recording quantity.

Quick check

1. Label each question as open or closed: (a) "Which of these do you eat most often: fruit / crisps / chocolate?" (b) "Describe the last time you felt nervous before a performance." (c) "On a scale of 1–7, how sleepy do you feel right now?" (d) "Why did you choose your current hobby?"
2. Dr Okafor assesses the same 30 children's motor skills (balance and coordination tasks) every 2 years, from age 6 until age 16. Name the research method, say how many times each child is tested, and give one problem she may face.
3. A researcher finds that the more hours students spend on part-time jobs, the lower their exam marks. Name the type of correlation and explain why this doesn't show that part-time jobs lower exam marks.
4. Lena joins a hiking club as a new member, without telling anyone she is a psychologist, and makes notes about how decisions are made in the group. Classify her observation as overt or covert and as participant or non-participant, giving a reason for each.
5. Arun thinks that people who walk around while learning a list of words will recall more of them than people who sit still, because earlier research found this. Write a suitable directional hypothesis and a null hypothesis.

Answers

1. (a) closed (fixed options); (b) open (own words); (c) closed (a rating scale is closed); (d) open.
2. A longitudinal study (the same children tested repeatedly over years). Tests at ages 6, 8, 10, 12, 14 and 16, so each child is tested **6** times. A problem: attrition, as families move away or change contact details, or children lose interest; the children may also improve on the tasks themselves through practice.
3. A **negative** correlation: as one co-variable (hours of work) goes up, the other (exam marks) goes down. A correlation can't show cause: students with lower marks might take more hours of work because they are less interested in study, or a third variable (such as needing to support their family) could cause both.
4. **Covert**, because the club members don't know she is observing them; **participant**, because she takes part in the club's hikes as a member.
5. Directional: "People who walk around a room while learning a list of 30 words will recall more words in a 2-minute written test than people who learn the list sitting at a desk." Null: "There will be no difference in the number of words recalled between people who walk while learning and people who sit while learning (any difference will be due to chance)."

Past questions to try

- 9990/22/F/M/25 Q1: overt observation and how it compares with covert.
- 9990/22/F/M/25 Q2: explaining what makes a study longitudinal.
- 9990/22/M/J/25 Q9: the two questionnaire techniques, applied to a scenario.
- 9990/22/O/N/24 Q4: repeated measures and counterbalancing in a lab experiment.