Experiments: laboratory and field

Lesson 2.1.1 | AS research methodology

Core ideas

Experiments investigate cause-and-effect by changing an independent variable and measuring a dependent variable.

Laboratory experiments take place in controlled settings, so procedures can be standardised and repeated more easily.

Field experiments take place in real-world settings, so behaviour may be more natural but harder to control.

Experimental and control conditions help researchers compare what changes when the independent variable is present or absent.

Random allocation, counterbalancing and clear instructions reduce bias and order effects.

Strong evaluation considers reliability, validity, ethics and whether findings transfer to real behaviour.

Experiments: laboratory and field

Lesson 2.1.1 | AS research methodology

Key terms

Experiment: A research method that changes one variable to see its effect on another.

Independent variable: The factor the researcher changes or compares.

Dependent variable: The behaviour or outcome the researcher measures.

Laboratory experiment: An experiment carried out in a controlled setting.

Field experiment: An experiment carried out in a real-world setting.

Control condition: A comparison condition without the key experimental change.

Random allocation: Putting participants into conditions by chance.

Ecological validity: How well a study reflects real-life behaviour.

Further detail and evidence

A field experiment still manipulates an IV; a naturalistic observation only records behaviour. Control means reducing alternative explanations, not merely using a laboratory. An active comparison or placebo can help distinguish the intended change from attention or expectations.

Experiments: laboratory and field

Lesson 2.1.1 | AS research methodology

Applying the ideas

Andrade randomly allocated adults to shape-shading or control conditions while keeping the recorded message constant. Better recall therefore supports the particular task comparison. Piliavin changed victim cues on a real train: realism increased, while passenger composition and actor differences were harder to control.

Worked example: Use the terms independent variable and dependent variable in a simple memory experiment. A researcher could change whether participants revise in silence or with background music; this is the independent variable. The researcher could then measure how many words each participant recalls from a list; this is the dependent variable. Comparing the recall scores helps test whether background music affected memory.

Worked example: Assess one strength and one weakness of laboratory experiments compared with field experiments. A strength of laboratory experiments is control: researchers can standardise instructions, reduce extraneous variables and repeat the procedure, which can improve reliability. A weakness is that the setting may feel artificial, so behaviour may not reflect real life as well as in a field experiment. This means laboratory experiments can gain control but lose ecological validity.

Evaluation and connections

Evaluate internal validity and ecological validity separately. A controlled task can identify an effect but transfer poorly to real life; a realistic field task can still contain confounds. Random allocation reduces systematic differences on average, but does not guarantee identical small groups. Covert field work can raise consent, withdrawal and debriefing concerns.

Experimental designs and controls

Lesson 2.1.2 | AS research methodology

Core ideas

Independent measures designs use different participants in each condition, which avoids order effects but can create participant-variable differences.

Repeated measures designs use the same participants in every condition, which controls participant variables but can create practice or fatigue effects.

Matched pairs designs pair similar participants before putting one member of each pair into each condition.

Order effects occur when doing one condition affects performance in another condition.

Counterbalancing changes the order of conditions across participants to reduce order effects.

Controls such as random allocation, standardised instructions and control groups make comparisons fairer.

Experimental designs and controls

Lesson 2.1.2 | AS research methodology

Key terms

Experimental design: The way participants are arranged across conditions.

Independent measures: Different participants are used in each condition.

Repeated measures: The same participants take part in every condition.

Matched pairs: Similar participants are paired before being split between conditions.

Order effects: Changes in performance caused by the order of conditions.

Counterbalancing: Varying the order of conditions to reduce order effects.

Random allocation: Putting participants into conditions by chance.

Extraneous variable: An unwanted factor that could affect the dependent variable.

Further detail and evidence

Independent measures uses different people in each condition; repeated measures uses the same people; matched pairs matches relevant characteristics and allocates pair members to different conditions. A control condition is the comparison treatment, not an experimental design in itself.

Experimental designs and controls

Lesson 2.1.2 | AS research methodology

Applying the ideas

For a silence/music memory study, repeated measures could use equivalent word lists and counterbalanced AB/BA orders. Different lists must be comparable. In matched pairs, match baseline recall before allocating one person to each condition. Bandura used matched triplets because there were three modelling conditions.

Worked example: Choose a suitable design for testing whether background music affects memory, and justify the choice. A repeated measures design could be suitable because the same participants could complete a memory task in silence and with background music. This controls participant variables such as natural memory ability. The researcher should counterbalance the order of the two conditions so practice or fatigue does not unfairly affect one condition.

Worked example: Assess one control that could make an experiment fairer. Random allocation can make an experiment fairer by placing participants into conditions by chance instead of by researcher preference. This reduces allocation bias and makes it less likely that one condition contains stronger or more motivated participants. It does not remove every extraneous variable, but it improves the fairness of the comparison.

Evaluation and connections

Counterbalancing distributes order effects rather than removing learning that cannot be reversed. Independent measures avoids carry-over but may need more participants. Matching reduces only measured differences and can be time-consuming; it does not make people identical. Explain which design best fits the actual task.

Self-reports: questionnaires and interviews

Lesson 2.1.3 | AS research methodology

Core ideas

Self-reports gather data directly from participants by asking them questions.

Questionnaires can collect data from many people efficiently, especially when questions are standardised.

Interviews allow follow-up questions and richer detail, but they take more time and may be affected by the interviewer.

Closed questions are easier to score and compare, while open questions can reveal more detailed answers.

Participants may give inaccurate answers because of social desirability, memory problems or misunderstanding questions.

Good self-report design uses clear wording, appropriate response options and ethical handling of personal information.

Self-reports: questionnaires and interviews

Lesson 2.1.3 | AS research methodology

Key terms

Self-report: A method where participants answer questions about themselves.

Questionnaire: A written or online set of questions completed by participants.

Interview: A spoken question-and-answer method.

Closed question: A question with fixed response options.

Open question: A question that lets participants answer in their own words.

Social desirability: Answering in a way that seems acceptable or impressive.

Response bias: A pattern of answers that reduces accuracy.

Standardised question: A question asked in the same way to every participant.

Further detail and evidence

Questionnaires may use paper-and-pencil or online techniques. Both can contain open and closed questions; technique and question format are different decisions.

A structured interview uses a fixed set/order of questions; an unstructured interview develops flexibly; a semi-structured interview uses prepared questions with follow-up prompts. Interviews may be face-to-face or by telephone.

Questionnaire technique (paper/online), question format (open/closed), interview format (structured/semi-structured/unstructured) and interview technique (face-to-face/telephone) are separate decisions. State example wording, response options and how answers will be scored or interpreted.

Self-reports: questionnaires and interviews

Lesson 2.1.3 | AS research methodology

Applying the ideas

A closed item might ask how many nights last week a person slept fewer than seven hours. An open item can ask what made sleep difficult. A structured interview asks the same prepared questions; a semi-structured interview can probe an unusual answer. Numerical scores and thematic descriptions can be combined.

Worked example: Choose either a questionnaire or an interview for researching exam stress, and justify the choice. A questionnaire would be suitable if the researcher wanted data from many students quickly. It could use closed rating-scale questions such as how often students feel stressed before exams, making answers easy to compare. A limitation is that students may underreport stress because of social desirability or because fixed options do not capture their full experience.

Worked example: Compare one strength of closed questions with one strength of open questions. Closed questions are useful because fixed options make answers quicker to complete and easier to score across many participants. Open questions are useful because participants can explain their views in their own words, giving richer qualitative detail. The best choice depends on whether the researcher values easy comparison or depth of response more.

Evaluation and connections

Clear wording avoids leading questions, double-barrelled items and unclear time periods. Pilot the questionnaire to check comprehension. Anonymous collection may reduce socially desirable responses but prevents some follow-up checks. Interviews allow clarification but may introduce tone, prompting or interviewer expectations.

Case studies

Lesson 2.1.4 | AS research methodology

Core ideas

Case studies focus on one individual, group or situation in much more detail than many other methods.

They often combine evidence from interviews, observations, tests, documents or background history.

Rich qualitative data can reveal unusual behaviour, rare conditions or complex real-life experiences.

Triangulation strengthens a case study by comparing evidence from more than one source or technique.

Case studies can have low generalisability because one case may not represent a wider population.

Evaluation should consider depth, validity, researcher bias, ethics, confidentiality and difficulty repeating the research.

Case studies

Lesson 2.1.4 | AS research methodology

Key terms

Case study: An in-depth investigation of one person, group, organisation or situation.

Qualitative data: Non-numerical data such as descriptions, explanations or interview comments.

Triangulation: Using more than one source or method to check evidence.

Generalisability: How far findings apply beyond the case studied.

Researcher bias: When a researcher's expectations influence interpretation.

Confidentiality: Protecting private information about participants.

Validity: How accurately research measures or represents what it claims to study.

Longitudinal evidence: Information collected over an extended period of time.

Further detail and evidence

A case study investigates one participant or defined unit in detail; the unit may be a person, family or organisation. It can combine interviews, observations, tests and documents. Triangulation compares different evidence, rather than simply collecting several copies of the same account.

Case studies

Lesson 2.1.4 | AS research methodology

Applying the ideas

Saavedra and Silverman combined the boy's history, child/parent interviews, observed button contact, distress ratings and follow-ups. Improved approach alongside initially rising distress showed why one measure could miss an important change. One case can generate a hypothesis for later controlled testing.

Worked example: Use a case-study method to investigate a rare phobia, and describe two sources of data. A psychologist could investigate one person with a rare phobia using interviews to gather the person's own account of fear, triggers and history. The psychologist could also use observations or rating scales during controlled exposure to record visible anxiety responses. Using more than one source would give richer evidence and support triangulation.

Worked example: Assess one strength and one limitation of case studies. A strength of case studies is depth because they can collect detailed qualitative information about a complex person or situation. A limitation is low generalisability because one case may not represent other people or contexts. This means case studies can be very informative, but conclusions should be made carefully.

Evaluation and connections

Rich contextual detail supports understanding of an unusual case, but researcher interpretation and selective recall can affect validity. One case need not represent a population. Detailed publication can reveal identity even without a name; protect privacy and obtain appropriate consent. Repeating the unique life history exactly is usually impossible.

Observations

Lesson 2.1.5 | AS research methodology

Core ideas

Observations collect data by watching and recording behaviour rather than asking participants to report it.

Naturalistic observations take place in ordinary settings, while controlled observations happen in arranged situations.

Overt observations are open to participants, while covert observations are hidden from them.

Participant observers join the group being studied; non-participant observers watch without joining in.

Behavioural categories help observers record behaviour consistently and reduce vague descriptions.

Inter-observer reliability improves when more than one observer records the same behaviour and their records agree.

Observations

Lesson 2.1.5 | AS research methodology

Key terms

Observation: A method where researchers watch and record behaviour.

Naturalistic observation: Observation in a real-world setting where behaviour normally occurs.

Controlled observation: Observation in a structured or arranged setting.

Overt observation: Observation where participants know they are being watched.

Covert observation: Observation where participants do not know they are being watched.

Participant observation: Observation where the researcher joins the group or activity.

Behavioural category: A clear behaviour label used when recording observations.

Inter-observer reliability: Agreement between observers recording the same behaviour.

Further detail and evidence

Structured observation uses predefined behavioural categories; unstructured observation records behaviour more openly. A category must describe observable behaviour clearly enough for different observers to apply it consistently.

Observation classifications can combine: an observation may be overt, non-participant, naturalistic and structured. Structured recording uses operationalised categories; unstructured recording keeps more open notes. The setting and the observer's participation are different dimensions.

Observations

Lesson 2.1.5 | AS research methodology

Applying the ideas

Define a queue-cutting objection as a spoken protest directed at the intruder, rather than a vague "negative reaction". Two observers independently code the same events, compare agreement and refine ambiguous categories. Video can allow rechecking, with appropriate consent and privacy protection.

Worked example: Choose one type of observation for studying playground helping behaviour, and justify the choice. A naturalistic observation would be suitable because helping behaviour could be watched in the playground setting where it normally happens. The researcher could use behavioural categories such as offering help, sharing equipment or calling an adult. This would give realistic data, although consent and privacy would need careful handling.

Worked example: Assess one way to improve reliability in an observational study. Inter-observer reliability can improve confidence in the data because two observers independently record the same behaviour and compare their records. If they agree often, the behavioural categories are probably clear and consistent. If they disagree, the categories or observer training may need to be improved.

Evaluation and connections

Observation measures visible behaviour rather than private motives. Overt observation can change behaviour; covert observation may reduce reactivity but compromises prior consent. Category overlap or missed events can reduce reliability. Agreement between observers does not establish that a category validly measures aggression, discomfort or another construct.

Correlations

Lesson 2.1.6 | AS research methodology

Core ideas

A correlation shows whether two measured variables are related.

A positive correlation means both variables tend to increase together.

A negative correlation means one variable tends to increase as the other decreases.

A correlation coefficient shows the direction and strength of the relationship, from -1 to +1.

Scattergraphs help students see patterns, outliers and whether the relationship is strong or weak.

Correlation does not prove causation because a third variable or reverse direction could explain the relationship.

Correlations

Lesson 2.1.6 | AS research methodology

Key terms

Correlation: A relationship between two measured variables.

Positive correlation: A relationship where both variables tend to increase together.

Negative correlation: A relationship where one variable increases as the other decreases.

No correlation: No clear relationship between the two variables.

Scattergraph: A graph showing paired scores for two variables.

Correlation coefficient: A number from -1 to +1 showing direction and strength.

Third variable: Another factor that may explain a relationship between two variables.

Causation: One factor directly producing a change in another.

Further detail and evidence

Co-variables are the two measured variables in a correlation. Operationalise each precisely, such as minutes of sleep recorded in a diary and number of words correctly recalled; neither variable is manipulated in a correlational study.

The correlation coefficient describes linear association: its sign gives direction and its absolute size gives strength. A coefficient near zero means little linear association, not proof that no possible relationship exists. Inspect a scatter graph for unusual points or non-linear patterns.

Correlations

Lesson 2.1.6 | AS research methodology

Applying the ideas

Longer REM periods accompanied longer dream reports in Dement and Kleitman. Both REM minutes and report word count were measured. Talkativeness or forgetting can influence report length. Do not relabel the measured co-variables as manipulated IV and DV.

Worked example: Interpret a correlation coefficient of +0.72 between revision time and test score. A coefficient of +0.72 suggests a fairly strong positive correlation. Students who revise for longer tend to have higher test scores, because the plus sign shows the variables move in the same direction. However, this does not prove revision caused the higher scores because other factors, such as prior knowledge or motivation, could also affect performance.

Worked example: Assess one limitation of using a correlation to study stress and sleep. A limitation is that a correlation may show that higher stress is connected with less sleep, but it cannot prove which variable caused the other. Stress might reduce sleep, poor sleep might increase stress, or a third variable such as workload could affect both. This makes correlations useful for spotting relationships but limited for explaining cause and effect.

Evaluation and connections

Correlations permit prediction and study variables that cannot ethically be manipulated. They do not establish direction of cause or eliminate third variables. A positive correlation is not automatically desirable, and a negative one is not automatically harmful. Keep the operational definitions attached to the relationship.

Longitudinal studies

Lesson 2.1.7 | AS research methodology

Core ideas

Longitudinal studies follow the same participants or groups across more than one time point.

They are useful for studying development, change, stability and long-term effects.

Repeated measurement can show patterns that a single snapshot study would miss.

Attrition happens when participants drop out, which can make later data less representative.

Repeated testing may change behaviour because participants become familiar with the task.

Evaluation should consider depth over time, cost, time, ethics, sample loss and whether findings still apply later.

Longitudinal studies

Lesson 2.1.7 | AS research methodology

Key terms

Longitudinal study: A study that collects data over an extended period of time.

Repeated measurement: Measuring the same participants or variables more than once.

Attrition: Loss of participants from a study over time.

Development: Change in behaviour, ability or experience over time.

Snapshot study: A study collecting data at one point in time.

Cohort effect: An influence linked to the particular generation or group studied.

Follow-up: A later stage of data collection after the first measurement.

Representativeness: How well a sample reflects the wider group it is meant to show.

Further detail and evidence

A longitudinal design repeatedly follows the same participants or unit. It can form part of an experiment, case study or observational study; "longitudinal" does not itself specify whether a variable was manipulated. State interval, task, scoring and re-contact arrangements.

Longitudinal studies

Lesson 2.1.7 | AS research methodology

Applying the ideas

Hölzel obtained structural scans before and after an eight-week course and used a comparison group. Saavedra and Silverman followed treatment with six- and twelve-month checks. Repeated measurement distinguishes immediate change from evidence that gains persist.

Worked example: Choose a longitudinal design for studying exam anxiety, and describe how data could be collected. A researcher could measure the same students' exam anxiety at the start of the course, before mock exams and before final exams. They could use the same anxiety questionnaire at each follow-up so scores can be compared across time. This would show whether anxiety changes as exams get closer, although some students might drop out before the final measurement.

Worked example: Assess one strength and one limitation of longitudinal studies. A strength is that longitudinal studies can show how behaviour changes over time, which a one-off snapshot study may miss. A limitation is attrition because participants may drop out, leaving a smaller or less representative sample. This means longitudinal evidence can be rich, but it can also be expensive, slow and affected by sample loss.

Evaluation and connections

Attrition threatens interpretation if dropouts differ from retained participants. Practice, changing instruments, life events and historical change can also influence results. Use suitable equivalent tasks and stable procedures where appropriate. A long follow-up can improve understanding while increasing cost and privacy burdens.

Aims and hypotheses

Lesson 2.2.1 | AS research methodology

Core ideas

An aim states the general purpose of a study in clear, student-friendly terms.

A hypothesis is a testable prediction about what the researcher expects to find.

A directional hypothesis predicts the direction of a difference or relationship.

A non-directional hypothesis predicts a difference or relationship without saying which way it will go.

A null hypothesis predicts no significant difference or relationship between the variables.

Good hypotheses use clear, measurable variables so the study can be tested and interpreted.

Key terms

Aim: The general purpose or question of a study.

Hypothesis: A testable prediction about the expected result.

Directional hypothesis: A prediction that states which way a difference or relationship will go.

Aims and hypotheses

Lesson 2.2.1 | AS research methodology

Continued

Non-directional hypothesis: A prediction that states there will be a difference or relationship but not its direction.

Null hypothesis: A prediction that there will be no significant difference or relationship.

Variable: A factor that can change or be measured.

Prediction: A statement about what is expected to happen.

Testable: Written so evidence can be collected to check it.

Further detail and evidence

A directional alternative hypothesis is also called one-tailed; a non-directional alternative hypothesis is also called two-tailed. Recognise both forms and the null hypothesis; statistical-test calculations are outside the 2027 syllabus.

An aim states what will be investigated. An alternative hypothesis predicts a difference or association; directional/one-tailed specifies its direction, while non-directional/two-tailed does not. A null predicts no population effect or relationship, with observed sample differences attributable to chance variation.

Aims and hypotheses

Lesson 2.2.1 | AS research methodology

Applying the ideas

Directional: adults listening in silence will recall more target words than adults hearing the specified background music. Non-directional: recall scores will differ between the conditions. Null: there will be no difference in recall scores between conditions beyond chance. Define the listening conditions and correct-word score clearly.

Worked example: Write a directional hypothesis for a study on sleep and memory. Participants who sleep for eight hours before a memory test will recall more words than participants who sleep for four hours. This is directional because it predicts which sleep group will have the higher memory score. It is testable because sleep duration and number of words recalled can be measured.

Worked example: Compare an aim with a hypothesis using a simple example. An aim could be to investigate whether background music affects concentration. A hypothesis is more specific, such as: students who work in silence will complete more concentration-task items correctly than students who work with background music. The aim gives the broad purpose, while the hypothesis predicts a measurable result.

Evaluation and connections

Use "relationship" for a correlation and "difference" for a comparison. Avoid vague predictions such as "music affects people". Hypotheses are general methods content and required in Paper 4 plans; the 2027 Paper 2 planning task does not require aims or hypotheses.

Independent variables and dependent variables

Lesson 2.2.2 | AS research methodology

Core ideas

The independent variable is the factor the researcher changes or compares.

The dependent variable is the outcome or behaviour the researcher measures.

In experiments, researchers look for whether changes in the independent variable affect the dependent variable.

The independent variable should have clear conditions, such as silence versus background music.

The dependent variable should be measurable, such as recall score, reaction time or rating-scale score.

Confusing the IV and DV can make aims, hypotheses and conclusions unclear.

Independent variables and dependent variables

Lesson 2.2.2 | AS research methodology

Key terms

Independent variable: The factor the researcher changes or compares.

Dependent variable: The behaviour or outcome the researcher measures.

Condition: A version or level of the independent variable.

Measure: A way of recording the dependent variable.

Experiment: A method that tests whether one variable affects another.

Outcome: The result recorded from participants.

Cause-and-effect: A claim that one factor produces a change in another.

Variable: A factor that can change or be measured.

Further detail and evidence

An IV needs identifiable conditions; a DV needs a score, count, time or other defined outcome. Naturally occurring participant differences are not randomly assigned manipulations. Correlational studies instead have two measured co-variables. Not every study has an IV.

Operationalise both variables: identify precisely what changes between conditions and how the outcome is scored. A clear label alone is not enough for another researcher to reproduce the manipulation or measurement.

Independent variables and dependent variables

Lesson 2.2.2 | AS research methodology

Applying the ideas

Andrade: IV = shading supplied shapes versus listening without shading; DV = adjusted recall of names/places. In a new study, "music" is too vague: define track, volume and duration. "Memory" is too vague: define the number of correct words recalled after a stated interval.

Worked example: Identify the IV and DV in a study testing whether caffeine affects reaction time. The independent variable is caffeine intake, such as no caffeine compared with a set dose of caffeine. The dependent variable is reaction time, measured by how quickly participants respond to a stimulus. This is clear because caffeine intake is the factor being compared and reaction time is the outcome recorded.

Worked example: Compare the role of the independent variable with the role of the dependent variable. The independent variable is what the researcher changes or compares to create conditions in a study. The dependent variable is what the researcher measures to see whether the independent variable made a difference. For example, background music could be the IV and number of words recalled could be the DV.

Evaluation and connections

Milgram's main 1963 condition measured maximum shock without manipulating a comparison IV. Naming an outcome does not turn any study into an experiment. A group difference supports a causal account only when design, allocation, controls and measurement justify it.

Operational definitions

Lesson 2.2.3 | AS research methodology

Core ideas

An operational definition states exactly how a variable is measured or manipulated in a study.

Vague concepts such as stress, memory or aggression need clear observable indicators.

Operationalising an independent variable means defining the exact conditions being compared.

Operationalising a dependent variable means defining the exact score, count or behaviour being recorded.

Clear operational definitions make studies easier to repeat and reduce disagreement between researchers.

A poor operational definition can weaken validity if it does not really capture the concept being studied.

Operational definitions

Lesson 2.2.3 | AS research methodology

Key terms

Operational definition: A precise statement of how a variable is measured or manipulated.

Operationalise: Turn a broad idea into a clear research measure or condition.

Observable behaviour: Behaviour that can be seen or recorded directly.

Measure: A specific way of recording data.

Manipulation: The exact change made to an independent variable.

Replication: Repeating a study to check whether similar results are found.

Validity: How well a measure captures what it claims to measure.

Reliability: How consistent a measure or procedure is.

Further detail and evidence

Operationalisation turns an abstract construct into a reproducible manipulation or measurement. Specify units, materials, timing, scoring and category boundaries. A precise measure can still lack validity if it captures only a poor substitute for the intended concept.

Operational definitions

Lesson 2.2.3 | AS research methodology

Applying the ideas

Operationalise helping as approaching and physically assisting within a stated observation period, recording latency in seconds. Operationalise questionnaire stress as the total on named items over a specified week. These measure different aspects of behaviour and experience; neither automatically captures everything relevant.

Worked example: Write an operational definition for stress in a study of students before exams. Stress could be operationalised as each student's score on a 20-item exam-anxiety questionnaire completed one week before the exam. This is measurable because every participant receives a numerical score using the same items. It is clearer than simply saying students are 'stressed' because researchers know exactly what data to record.

Worked example: Assess why operational definitions matter when studying aggression. Operational definitions matter because aggression is a broad idea that different observers might interpret differently. If researchers define aggression as the number of hits, kicks and shouted insults in a set time, observers can record the same behaviours more consistently. A limitation is that this definition may miss less visible forms of aggression, so validity still needs checking.

Evaluation and connections

A broad aggression category may be ambiguous; an overly narrow one may miss important acts. Pilot categories with independent observers. Word count is a reproducible report-length measure but not exact dream duration. Evaluate the link between the indicator and the construct as well as consistency.

Control of variables and standardisation

Lesson 2.2.4 | AS research methodology

Core ideas

Control of variables means reducing unwanted factors that could affect the dependent variable.

Extraneous variables are possible influences that are not the main focus of the study.

Standardisation means keeping instructions, materials and procedures the same for all participants where possible.

Control groups or comparison conditions help researchers judge whether the independent variable made a difference.

Random allocation can reduce bias by assigning participants to conditions by chance.

Strong evaluation balances control and reliability against artificiality and ecological validity.

Control of variables and standardisation

Lesson 2.2.4 | AS research methodology

Key terms

Control of variables: Managing unwanted factors so comparisons are fairer.

Standardisation: Keeping procedures the same across participants or conditions.

Extraneous variable: An unwanted factor that could affect the dependent variable.

Control group: A comparison group that does not receive the key experimental treatment.

Comparison condition: A condition used to compare against another condition.

Random allocation: Putting participants into conditions by chance.

Procedure: The planned steps followed in a study.

Ecological validity: How well a study reflects real-life behaviour.

Further detail and evidence

Participant variables include abilities and previous experience; situational variables include noise, instructions and time. An extraneous variable becomes a confound when it systematically varies with the IV and offers an alternative explanation of the DV difference.

Control of variables and standardisation

Lesson 2.2.4 | AS research methodology

Applying the ideas

Keep a memory test's word list, exposure time, delay and instructions constant, while changing only the defined music condition. Random allocation distributes participant differences. Counterbalancing distributes order; matched pairs addresses selected baseline characteristics. A placebo can control expectations about receiving treatment.

Worked example: Choose two controls for a memory experiment and explain why they matter. A researcher could give every participant the same word list and the same amount of time to learn it. These controls matter because differences in word difficulty or learning time could affect recall scores instead of the independent variable. Keeping these parts standardised makes the comparison fairer.

Worked example: Assess one strength and one limitation of standardising a procedure. A strength of standardisation is that it improves reliability because the study can be repeated using the same instructions, materials and timing. A limitation is that a highly standardised setting may feel artificial, so behaviour may not fully match real life. Standardisation helps consistency, but ecological validity still needs evaluation.

Evaluation and connections

Standardisation applies to relevant procedures, not to making the intended IV identical. Excessive control can create artificiality. Blinding reduces some expectation effects, but may be impossible when a treatment is obvious. Name a plausible unwanted influence and explain how the proposed control reduces it.

Quantitative and qualitative data

Lesson 2.3.1 | AS research methodology

Core ideas

Quantitative data are numerical, such as scores, counts, times, percentages or ratings.

Qualitative data are descriptive, such as interview comments, written explanations or observation notes.

Quantitative data are useful for comparisons, summaries, graphs and statistical analysis.

Qualitative data can give richer detail about meanings, experiences and reasons for behaviour.

Some studies collect both types of data to balance measurement with depth.

Evaluation should consider objectivity, detail, ease of analysis, interpretation and whether the data answer the research question.

Quantitative and qualitative data

Lesson 2.3.1 | AS research methodology

Key terms

Quantitative data: Numerical data such as scores, counts or times.

Qualitative data: Descriptive data such as words, explanations or observations.

Numerical score: A number used to record a response or performance.

Descriptive detail: Information that explains behaviour or experience in words.

Rating scale: A scale that turns a judgement into a number.

Statistical analysis: Using numerical methods to summarise or compare data.

Theme: A repeated idea or pattern found in qualitative data.

Interpretation: Making meaning from data after it is collected.

Further detail and evidence

Quantitative and qualitative describe the form of data, not whether it is true or unbiased. A numerical distress rating remains subjective. Qualitative observation can use systematic procedures. Converting comments into counts loses some context while allowing comparisons.

Quantitative and qualitative data

Lesson 2.3.1 | AS research methodology

Applying the ideas

Piliavin recorded response times and counts, alongside passenger comments. Saavedra and Silverman combined a 0-8 distress score with descriptions and observation. The methods complement one another: an average can reveal a pattern while detailed accounts suggest why or how it occurred.

Worked example: Classify two pieces of data from a memory study as quantitative or qualitative. A score of 16 words recalled out of 20 is quantitative because it is numerical and can be compared across participants. A participant's comment that the word list felt confusing because the words were unrelated is qualitative because it gives descriptive detail in words. The two types of data answer different questions about performance and experience.

Worked example: Assess one strength of using qualitative data alongside quantitative data. Using qualitative data alongside quantitative data can add depth to numerical results. For example, test scores may show that one group performed worse, while interview comments may suggest they felt distracted or anxious. This combination can make interpretation richer, although qualitative data can take longer to analyse and may be more subjective.

Evaluation and connections

Numerical data are often economical to summarise, but can hide variation or meaning. Qualitative material provides detail but takes time and involves interpretation. Mixed data do not automatically guarantee validity; explain how each measure answers the actual research question.

Subjective and objective data

Lesson 2.3.2 | AS research methodology

Core ideas

Subjective data depend on personal opinion, interpretation or judgement.

Objective data are recorded in a way that is less dependent on personal judgement.

Self-reports often produce subjective data because participants describe their own feelings or experiences.

Objective measures can include reaction times, recall scores, counts of behaviour or physiological readings.

Subjective data can give insight into personal experience, but may be affected by bias or social desirability.

Objective data can improve reliability, but may miss meanings or feelings that are important to the research question.

Subjective and objective data

Lesson 2.3.2 | AS research methodology

Key terms

Subjective data: Data influenced by personal opinion, judgement or interpretation.

Objective data: Data recorded with less dependence on personal judgement.

Self-report: A method where participants answer questions about themselves.

Rating scale: A numerical scale used to record a judgement or feeling.

Behaviour count: Recording how often a specific behaviour occurs.

Physiological measure: A body-based measure such as heart rate or brain activity.

Bias: A factor that can distort data or interpretation.

Interpretation: Making meaning from data after it is collected.

Further detail and evidence

Objective recording is less dependent on personal judgement; subjective data express an experience or interpretation. This distinction differs from numerical/descriptive form. Instrument readings can be objective while their interpretation as a psychological state remains uncertain.

Subjective and objective data

Lesson 2.3.2 | AS research methodology

Applying the ideas

An eye electrode records a physiological signal; the dream report describes subjective experience. A numerical pain score is quantitative but subjective. A bottle-opening time is objectively recorded, but opening does not prove medication ingestion. Identify the measure and the construct separately.

Worked example: Classify two data examples from a stress study as subjective or objective. A student's rating of how stressed they feel from 1 to 10 is subjective because it depends on the student's personal judgement. A heart-rate reading during the same task is more objective because it is recorded by equipment and is less dependent on opinion. Both can be useful because they show different sides of stress.

Worked example: Assess one reason to collect both subjective and objective data. Collecting both types can give a fuller picture. For example, a participant may show a high heart rate during a task, but their self-report can reveal whether they actually felt anxious or excited. The objective data can support reliability, while the subjective data adds meaning, although the two may not always agree.

Evaluation and connections

Subjective information can be essential when the construct is pain or feeling distressed. Objective indicators may improve reproducibility without capturing personal meaning. Neither type is inherently perfect; assess calibration, reporting bias, inference and fitness for the question.

Samples, populations and sampling techniques

Lesson 2.3.3 | AS research methodology

Core ideas

A population is the wider group the researcher wants to understand.

A sample is the smaller group of participants who actually take part in the study.

Representative samples make it easier to generalise findings to the wider population.

Opportunity sampling uses people who are available, making it practical but often biased.

Random sampling gives each member of the population an equal chance of selection if a full sampling frame is available.

Sampling technique affects bias, ethics, time, cost and how confidently findings can be applied beyond the study.

Key terms

Population: The wider group the researcher wants to draw conclusions about.

Sample: The participants who actually take part in a study.

Sampling technique: The method used to choose participants.

Representative sample: A sample that reflects important features of the wider population.

Opportunity sample: A sample chosen because participants are available.

Samples, populations and sampling techniques

Lesson 2.3.3 | AS research methodology

Continued

Random sample: A sample where each population member has an equal chance of selection.

Sampling bias: A sample problem that makes some people more likely to be included than others.

Generalisability: How far findings apply beyond the sample studied.

Volunteer sampling: Recruiting participants who choose to respond to an invitation; also called self-selecting sampling.

Further detail and evidence

Volunteer (self-selecting) sampling recruits people who respond to an invitation or advertisement. It is practical and participants may be motivated, but volunteers may differ from non-volunteers, limiting generalisability.

Random sampling selects from the target population by chance. Random allocation instead assigns already-recruited participants to experimental conditions; these are separate procedures.

Define the target population and the sampling frame. Random sampling selects people from that frame; opportunity sampling recruits available people; volunteer/self-selection sampling uses responses to an invitation. Random allocation happens later and assigns participants to conditions.

Samples, populations and sampling techniques

Lesson 2.3.3 | AS research methodology

Applying the ideas

Milgram recruited volunteers through advertisements/direct mail. Andrade recruited available panel members after another experiment, then randomly allocated them. A school sample may be practical but exclude children who do not attend those schools. Large numbers alone do not remove selection bias.

Worked example: Choose a sampling technique for a quick school stress survey and justify it. Opportunity sampling would be practical for a quick school stress survey because the researcher could ask students who are available at lunch or in a common area. This would save time and make data collection easier. A limitation is sampling bias because available students may not represent all students in the school.

Worked example: Compare a sample with a population using a psychology example. If a researcher wants to understand exam anxiety in all Year 12 students at a school, that full group is the population. If 40 Year 12 students complete the questionnaire, those 40 students are the sample. The sample is easier to study, but conclusions are stronger if it represents the wider population well.

Evaluation and connections

A random sample can still be affected by nonresponse or an incomplete frame. Volunteers may differ in motivation and relevant traits. Generalisation depends on the sample, target population, setting and task. Describe the actual recruitment information rather than assuming a technique not reported.

Measures of central tendency and spread

Lesson 2.3.4 | AS research methodology

Core ideas

Measures of central tendency describe a typical score in a set of data.

The mean is calculated by adding all scores and dividing by the number of scores.

The median is the middle score when scores are placed in order.

The mode is the most frequent score in a data set.

The range shows spread by subtracting the lowest score from the highest score.

Outliers can distort the mean and range, so the median may sometimes be a better summary.

Key terms

Central tendency: A summary of the typical or middle value in a data set.

Mean: The total of all scores divided by the number of scores.

Median: The middle score when values are ordered.

Mode: The most common score in a data set.

Range: The highest score minus the lowest score.

Measures of central tendency and spread

Lesson 2.3.4 | AS research methodology

Continued

Outlier: An unusual score that is far from most other scores.

Data set: A collection of scores or values.

Spread: How far apart or varied the scores are.

Standard deviation: A measure of the dispersion of scores around their mean; larger values indicate greater variability.

Further detail and evidence

Standard deviation describes how dispersed scores are around the mean: a smaller standard deviation means scores are more closely clustered; a larger one means greater variability. Compare it with the mean rather than treating it as a measure of the average.

Two groups can have the same mean but different standard deviations. If both mean recall scores are 10, SD 1 indicates more consistent scores than SD 4; it does not mean the first group has higher average recall.

Mean summarises all scores; median is the middle ordered score; mode is most frequent. Range is the highest minus lowest score. Standard deviation describes dispersion around the mean: smaller values indicate tighter clustering. Know how these measures are found and interpret them.

Measures of central tendency and spread

Lesson 2.3.4 | AS research methodology

Applying the ideas

Two conditions can both have a mean recall score of 10, while SD 1 versus SD 4 indicates different consistency, not a difference in average recall. A very high outlier can raise the mean and range while having less effect on the median. Compare groups meaningfully, keeping units clear.

Worked example: Calculate the mean and range for the scores 4, 6, 6, 8 and 11. The mean is 7 because the total is 35 and there are 5 scores, so 35 divided by 5 equals 7. The range is 7 because the highest score is 11 and the lowest score is 4, so 11 minus 4 equals 7. These two measures summarise a typical score and the spread of scores.

Worked example: Assess why a researcher might report the median instead of the mean. A researcher might report the median when a data set contains an outlier. For example, if most reaction times are close together but one participant is extremely slow, the mean could be pulled upward and become less representative. The median is less affected by the extreme score, so it may better show the typical performance.

Evaluation and connections

A mean without spread can conceal important variation. The mode suits the most common category; a mean is not meaningful for arbitrary category labels. The 2027 syllabus does not require candidates to perform calculations or interpret unspecified inferential tests. Optional arithmetic practice must not be presented as an extra exam requirement.

Tables, bar charts, histograms and scatter graphs

Lesson 2.3.5 | AS research methodology

Core ideas

Tables organise data in rows and columns so values can be read accurately.

Bar charts compare separate categories, such as two conditions or participant groups.

Histograms show the distribution of continuous data using touching bars.

Scatter graphs show the relationship between two measured variables.

Graphs should have clear titles, labelled axes and suitable scales.

Choosing the wrong graph can make patterns harder to interpret or misrepresent the data.

Key terms

Table: A display of data arranged in rows and columns.

Bar chart: A graph with separate bars used to compare categories.

Histogram: A graph with touching bars used for continuous data intervals.

Scatter graph: A graph showing paired scores for two variables.

Axis: A labelled line on a graph showing what is measured.

Scale: The numbered intervals used on an axis.

Category: A separate group or condition in the data.

Continuous data: Numerical data that can take values along a scale.

Tables, bar charts, histograms and scatter graphs

Lesson 2.3.5 | AS research methodology

Further detail and evidence

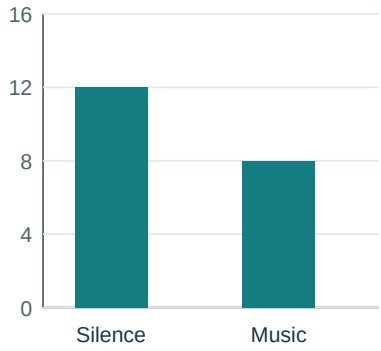
2027 graph skills include naming, recognising, drawing, changing and interpreting bar charts, histograms and scatter graphs, and presenting/interpreting tables. Reading a graph alone does not cover the drawing and representation requirements.

Bar charts compare categories with separated bars. Histograms show grouped continuous scores with adjoining bars; equal-width intervals are simplest to interpret. A scatter graph plots one point per paired observation. A table needs clear headings, units and totals where relevant.

Choosing a graph

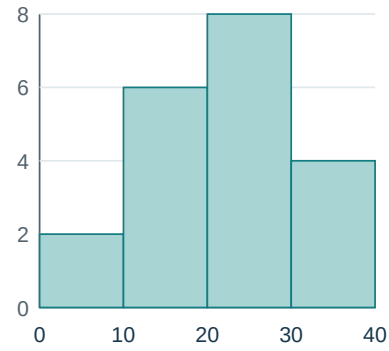
Illustrative data: these figures are not results from a named study.

Bar chart



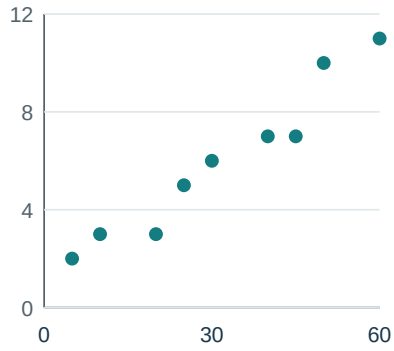
Mean recalled words

Histogram



Response time (s)
Vertical axis: frequency

Scatter graph



Study time (min)
Vertical axis: recall

Tables, bar charts, histograms and scatter graphs

Lesson 2.3.5 | AS research methodology

Applying the ideas

A bar chart could compare mean recall in silence/music conditions. A histogram could show the distribution of response times. A scatter graph could plot each person's sleep minutes against recall score. Choose the form from the research question and data, rather than treating every graph as interchangeable.

Worked example: Choose the best graph for showing the relationship between revision time and test score, and justify it. A scatter graph would be best because revision time and test score are two measured variables for each participant. Each point could show one student's revision time and score, making it possible to see whether there is a positive, negative or no correlation. A bar chart would be less suitable because the aim is to show a relationship, not compare separate categories.

Worked example: Assess why clear graph labels and scales matter in psychological research. Clear labels and scales matter because readers need to know what each axis shows and how the values are spaced. If a scale is misleading or an axis is not labelled, the pattern may be misread, such as exaggerating a small difference between conditions. Good presentation makes the data easier to interpret accurately.

Evaluation and connections

Graph skills include recognising, drawing, changing representations and interpreting. Label axes and units, use an appropriate consistent scale, and keep category gaps distinct from continuous intervals. A truncated or inconsistent scale can exaggerate a difference. A scatter trend supports association, not causation.

Human ethical guidelines

Lesson 2.4.1 | AS research methodology

Core ideas

Ethical guidelines help researchers treat human participants responsibly and respectfully.

Informed consent means participants should understand what they are agreeing to before taking part.

The right to withdraw means participants can leave a study without pressure or penalty.

Protection from harm means researchers should avoid unnecessary physical or psychological risk.

Confidentiality means personal information should be protected and not reported in an identifiable way.

Debriefing explains the study afterwards, especially if deception was used or participants may feel concerned.

Human ethical guidelines

Lesson 2.4.1 | AS research methodology

Key terms

Ethical guideline: A principle for conducting research responsibly.

Informed consent: Agreement to take part after understanding the study.

Right to withdraw: The right to leave a study without pressure or penalty.

Protection from harm: Avoiding unnecessary physical or psychological risk.

Confidentiality: Protecting personal information from being identified.

Deception: Misleading participants or withholding the full aim.

Debriefing: Explaining the study and addressing concerns afterwards.

Privacy: Respecting a person's personal space, information and choices.

Further detail and evidence

Human ethics includes valid informed consent, right to withdraw, minimising harm and maximising benefit, lack of deception, confidentiality, privacy and debriefing. Privacy concerns access to a person or their behaviour; confidentiality concerns protection of identifiable information.

Privacy concerns access to the person, behaviour or personal experiences; confidentiality concerns protecting identifiable data. Avoid deception where possible, explain participation accurately, minimise harm while maximising benefit, maintain withdrawal, and give an appropriate debrief.

Human ethical guidelines

Lesson 2.4.1 | AS research methodology

Applying the ideas

Milgram involved concealed aims and apparent shocks, visible distress and pressure from prods. Pozzulo documented parental consent, children's freedom to stop and distress monitoring. These examples support specific ethical analysis instead of assuming every laboratory study has the same risks.

Worked example: For a study using a stressful memory task, name two ethical safeguards and say how each would work. Researchers should protect participants from harm by stopping the task if someone becomes distressed and by avoiding unnecessarily upsetting materials. They should also give the right to withdraw by telling participants they can leave at any point without penalty. These steps reduce risk while still allowing the memory task to be studied.

Worked example: Why should participants be debriefed after a study that used deception? Deception can mean participants did not fully understand the true aim or procedure when they agreed to take part. A debriefing can explain what really happened, why deception was used, and give participants a chance to ask questions or withdraw their data. This does not make deception automatically acceptable, but it helps reduce harm and restore informed understanding.

Evaluation and connections

Consent is ongoing, and a signature does not remove later withdrawal. Children need suitable guardian permission and willing assent. Debriefing does not automatically undo harm. Distinguish documented safeguards, plausible risks and details not reported; absence of reporting is not proof that every safeguard was absent.

Animal ethical guidelines

Lesson 2.4.2 | AS research methodology

Core ideas

Animal research must balance possible scientific or practical value against harm to the animals involved.

Replacement means using a non-animal method where a suitable alternative exists.

Reduction means using the fewest animals needed to answer the research question properly.

Refinement means improving housing, handling and procedures to reduce pain, stress or restriction.

Researchers should consider species-specific needs, including social behaviour, diet, space and enrichment.

Ethical evaluation should include welfare, consent limitations, procedure severity, long-term effects and whether the findings justify the cost.

Animal ethical guidelines

Lesson 2.4.2 | AS research methodology

Key terms

Animal welfare: The physical and psychological well-being of an animal.

Replacement: Using non-animal alternatives where possible.

Reduction: Using the fewest animals needed for useful evidence.

Refinement: Changing procedures or care to reduce harm and improve welfare.

Species-specific need: A requirement linked to the natural behaviour or biology of a species.

Enrichment: Features that support natural behaviour and stimulation.

Procedure severity: How much pain, stress or restriction a procedure may cause.

Scientific value: The usefulness of research for knowledge or practical benefit.

Further detail and evidence

The 2027 animal-ethics checklist explicitly includes replacement, suitability of species, the number of animals, and procedures covering pain/suffering/distress, housing, rewards, deprivation and aversive stimuli. Judge these against the scientific value and minimisation of harm.

The official checklist includes harm/benefit, replacement, species, numbers and procedures. Procedures include pain/suffering/distress, housing, rewards, deprivation and aversive stimuli. Reduction/refinement are useful supporting ideas, but do not replace the explicit checklist.

Animal ethical guidelines

Lesson 2.4.2 | AS research methodology

Applying the ideas

Hassett maintained group housing and normal provisioning, but toy damage required monitoring. Fagen used rewards without adding aversive training in sessions, while pre-existing chains restricted movement. An animal's ability to turn away matters, but is not informed human consent.

Worked example: A researcher wants to study learning in animals. Which two of the 3Rs should shape the plan? Replacement should be considered first, because the researcher should ask whether a simulation, existing data or non-animal method could answer the question. Reduction should also shape the plan by using only the number of animals needed for reliable evidence. Together, these principles reduce unnecessary animal use before the study begins.

Worked example: How could refinement improve welfare in an animal-learning study? Refinement could involve using positive reinforcement instead of punishment, providing suitable housing and enrichment, and keeping procedures short enough to reduce stress. The researcher should also monitor animals for signs of distress and stop or adjust the procedure if welfare is affected. This keeps the research more humane while still allowing behaviour to be studied.

Evaluation and connections

Justify species and sample size for a worthwhile question. Too few animals can produce unusable evidence; unnecessary animals increase welfare cost. Existing husbandry remains relevant even when the study did not introduce it. Potential veterinary benefit is different from a measured welfare improvement.

Validity and ecological validity

Lesson 2.4.3 | AS research methodology

Core ideas

Validity asks whether a study or measure accurately represents the behaviour or concept being investigated.

Ecological validity asks whether findings reflect real-life behaviour outside the research setting.

A task can be well controlled but still lack validity if it does not measure the intended behaviour well.

Artificial materials, unusual settings and demand characteristics can reduce ecological validity.

High ecological validity can make findings more realistic, but may reduce control over extraneous variables.

Strong evaluation explains the specific validity issue rather than only saying a study is artificial.

Validity and ecological validity

Lesson 2.4.3 | AS research methodology

Key terms

Validity: How well research measures or represents what it claims to study.

Ecological validity: How well findings reflect real-life behaviour.

Task validity: How well a task measures the behaviour or ability it is meant to measure.

Demand characteristics: Clues that may make participants guess the aim and change behaviour.

Artificial setting: A research setting that differs from everyday life.

Realism: How closely a study resembles normal experience or behaviour.

Extraneous variable: An unwanted factor that could affect the result.

Generalisability: How far findings apply beyond the sample or setting studied.

Further detail and evidence

Validity concerns whether a measure or conclusion captures the intended concept. Ecological validity concerns relevance to everyday behaviour. Demand characteristics can alter responses when participants infer expectations. Generalisability extends to populations and contexts, not only the physical setting.

Validity and ecological validity

Lesson 2.4.3 | AS research methodology

Applying the ideas

Bobo attacks may represent invited toy play rather than ordinary violence. A computerised personal-space choice is controlled but indirect. A line-up silhouette permits rejection, yet benign cartoons differ from a stressful crime. Identify the construct and the exact task before explaining the limitation.

Worked example: A memory study uses a list of random words. What validity issue could this create? A random word list may reduce ecological validity because real memory often involves meaningful information, conversations, routes or events rather than isolated words. The task may still be controlled and easy to score, but it may not fully represent everyday memory. The researcher should be careful when applying the findings to real-life remembering.

Worked example: Why might a field study improve ecological validity but reduce control? A field study may improve ecological validity because behaviour is observed in a more natural setting where people behave normally. However, the researcher may have less control over extraneous variables such as noise, crowding or other people's actions. This means the findings may be more realistic but harder to interpret as clear cause and effect.

Evaluation and connections

A realistic field task can have confounds; a laboratory task can still investigate a meaningful process. Objectivity does not ensure construct validity, and reliability does not ensure correctness. Evaluate the specific link between procedure, score and intended claim.

Reliability and replicability

Lesson 2.4.4 | AS research methodology

Core ideas

Reliability is about consistency in procedures, measures or findings.

A reliable measure gives similar results when used in the same way under similar conditions.

Inter-observer reliability checks whether different observers record behaviour in similar ways.

Test-retest reliability checks whether a measure gives similar results across time when the target should be stable.

Replicability means a study can be repeated to see whether similar findings are found again.

High reliability supports confidence in evidence, but a reliable measure still needs to be valid.

Key terms

Reliability: Consistency of a measure, procedure or finding.

Replicability: The ability to repeat a study and check whether similar findings occur.

Inter-observer reliability: Agreement between observers recording the same behaviour.

Reliability and replicability

Lesson 2.4.4 | AS research methodology

Continued

Test-retest reliability: Consistency of scores when the same measure is used again later.

Standardised procedure: A procedure kept the same across participants or repeats.

Consistency: Similarity across measurements, observers or occasions.

Measurement error: Inaccuracy or variation caused by the measuring process.

Validity: How well a study or measure captures what it claims to study.

Further detail and evidence

Inter-rater reliability is agreement between people judging or scoring the same material; inter-observer reliability applies to observers recording behaviour. Describe how agreement would be checked and improve unclear categories if agreement is poor.

Inter-rater reliability is agreement between people scoring the same material; inter-observer reliability concerns behaviour recording. Test-retest reliability assesses consistency across time when the target should be stable. Replicability requires enough procedural detail for another researcher to repeat the study.

Reliability and replicability

Lesson 2.4.4 | AS research methodology

Applying the ideas

Train independent observers using the same category definitions and compare coding of the same events. If agreement is poor, identify ambiguous categories and recheck after refinement. For a stable questionnaire construct, repeat the same measure after a suitable interval while considering memory effects.

Worked example: A playground observation uses two observers. How could they check reliability? The observers could independently record the same playground behaviour using the same behavioural categories, then compare how often their records match. High agreement would suggest good inter-observer reliability. If agreement is low, the categories may need clearer definitions or the observers may need more training.

Worked example: Why is a replicable study more useful than a one-off finding? A replicable study is more useful because other researchers can repeat the procedure and check whether similar results occur again. If findings appear across repeated studies, psychologists can have more confidence that the result was not just chance, bias or a one-time mistake. Replicability strengthens evidence, although the original measure still needs to be valid.

Evaluation and connections

Agreement does not establish validity: two observers may share the same flawed interpretation. Change across time can be genuine rather than unreliability. Standardised instructions and scoring support consistency; reporting only "it is reliable because it is a lab" does not explain a check or its result.

Planning studies for Paper 2

Lesson 2.4.5 | AS research methodology

Core ideas

Paper 2 Section B contains a planning question worth 14 marks in total. Use the information given and plan the missing parts of an investigation, then evaluate practical/methodological issues, validity and reliability.

Include the procedure, how the plan supports validity and reliability, and the type of data. Give method-specific details that make the proposed investigation usable, rather than listing definitions.

For experiments, specify the operationalised IV and DV, controls/standardisation and experimental design; for a field experiment also specify the location. For self-reports, specify question format, example questions, scoring/interpretation, and questionnaire technique or interview format and technique.

For a case study, identify the participant/unit, information collected, at least two data-collection techniques and analysis/interpretation/triangulation. For observation, specify overt/covert, participant/non-participant, naturalistic/controlled and structured behavioural categories or unstructured recording.

For correlation, specify two co-variables, how each is measured and the proposed relationship/scatter graph. For a longitudinal study, specify tests/tasks, scoring, frequency/interval, how participants are re-contacted and controls/standardisation.

Paper 2 scope: aims and hypotheses are not expected in the plan. Candidates are not expected to evaluate sample/generalisability or ethics of the plan. Ethics remains required elsewhere in the course; these limits apply specifically to this planning task.

Planning studies for Paper 2

Lesson 2.4.5 | AS research methodology

Key terms

Planning: Describing a usable proposed investigation and explaining the choices relevant to the question.

Triangulation: Comparing information from at least two sources or techniques to check an interpretation.

Further detail and evidence

For the 2027 Paper 2 plan, give procedure, validity, reliability and data type, plus the required details of the named method. Experiments need defined IV/DV, controls and design; field experiments also need location. Correlations need both co-variables, their measures and the proposed relationship/scatter graph.

Applying the ideas

A memory plan can specify a fixed word list, timed exposure, defined delay and a correct-word recall score. State the music/silence conditions and how allocation/design works. Then explain how equal materials reduce a named confound and how standard scoring supports consistency. Follow any features supplied in the question.

Planning studies for Paper 2

Lesson 2.4.5 | AS research methodology

Continued

Worked example: Plan a naturalistic observation of helping in a school common area. Specify what will be recorded and explain one way to improve reliability. Use non-participant observation in the normal common area during a fixed ten-minute period. Define helping as picking up and returning a dropped object, and count each completed act using a structured checklist. Use two observers trained on the same definition to record the same period independently, then compare their records and clarify categories where they disagree. Keep observation periods and the checklist consistent.

Worked example: Plan a longitudinal self-report study of exam stress over a school term. Give the timing, measure, scoring and re-contacting procedure. Use the same short questionnaire at the start, midpoint and end of term. Include an identical item such as "During the past seven days, how often have you felt worried about school tests?" with ordered responses from never to very often, scored consistently from 0 to 4. Assign a participant code and use agreed reminders to re-contact the same participants at each interval. Keep the wording, reference period and instructions unchanged, and compare scores for the same codes over time.

Evaluation and connections

Do not add aims/hypotheses as expected Paper 2 plan features. The specific planning evaluation does not require sample/generalisability or ethics, although these remain examinable methods content elsewhere. Case studies need at least two collection techniques and triangulation; longitudinal plans need interval, scoring and re-contact arrangements.

Hint: follow the method and supplied details in the question. Write the missing procedure precisely and connect each proposed control to validity or reliability; a generic claim that a plan is reliable earns less useful explanation.